



CPIPC

中国研究生
创新实践系列大赛



中国研究生创新实践系列大赛
“华为杯”第十六届中国研究生
数学建模竞赛

学 校

上海海事大学

参赛队号

19102540248

1. 王 鑫

队员姓名

2. 曾叙砜

3. 唐 雁

中国研究生创新实践系列大赛
“华为杯”第十六届中国研究生
数学建模竞赛

题 目

全球变暖？

摘要：

本文针对人们所关心的全球变暖问题，展开了一系列的讨论：

针对问题一，首先考虑挖掘能够代表加拿大各地区温度变化的气温数据，经过数据挖掘、清洗等处理过程，得到加拿大 13 个地区 2000-2019 年的气温数据。通过描述性统计分析、回归分析、正态拟合、M-K 突变检验等统计分析方法对加拿大气温变化的时空分布及变化趋势。第二小问，首先通过提取特定地点的温度信息，分析加拿大西南海域的温度变化特征，接着通过 EOF 分解，得到全球 SSTA（海洋表面温度异常场）的五个主成分（模态），并依次对五个模态进行分析，得到全球海洋表面温度变化异常的主要变化规律。

针对问题二，首先对现有资料进行分析，寻找影响气候变化的要素，为后期气候预测提取相关数据做准备；在分析对比既有预测方法优缺点的基础上，探究满足全球气候特征数据的预测方法，具体针对数据特点，建立精度高、鲁棒性、自适应高的预测模型，即具有鲁棒性的回声状态网络预测方法；通过对影响因素的量化，带入模型对全球温度变化进行预测，其中通过对比模拟预测值与实际观测值验证了模型的有效性；最后，同比其他预测模型，证明该模型的优势。

针对问题三，首先从极寒天气发生地区的地理特征出发，分析极寒天气的影响因素，从而简单论证了极寒天气与气候变化之间的关系。然后，我们借助极值理论，并选取加拿大 NUNAVUT 地区为例，分析该地区近 20 年来极寒天气出现的变化特征，初步讨论某地极寒天气的出现与该地区气候变化的关系。然后从全球性的角度，通过构建分层的贝叶斯时空模型，进一步分析全球变暖和局地极寒现象的出现之间是否矛盾。

针对问题四，首先我们运用简单阐述了“全球变暖了，某地今年的冬天特别冷”之间的关系。紧接着，结合系统动力学的分析方法，重新介绍了“全球变暖”的概念，并且考虑了大气系统内部各圈层复杂的相互作用，以及两个重要的大气系统外界因素对气候变化的影响，对气候变化的趋势以及复杂性进行分析。

关键词：EOF 气候预测 非线性时间序列 分层贝叶斯时空模型 系统动力学

目录

1. 问题重述	4
1.1 问题的背景	4
1.2 要解决的问题	4
2. 模型假设与前期任务	5
3. 问题一：气候数据挖掘	6
3.1 加拿大各地温度的时空变化趋势	6
3.1.1 加拿大概况	6
3.1.2 研究方法	7
3.1.3 数据处理	8
3.1.4 分析结果	9
3.2 海洋表面温度数据规律	16
3.2.1 研究背景	16
3.2.2 相关理论	17
3.2.3 研究方法	17
3.2.4 数据处理	18
3.2.5 分析结果	18
3.3 本章小结	24
4. 问题二：气候变化预测	25
4.1 气候预测研究综述	25
4.1.1 研究现状	25
4.1.2 研究方法	27
4.1.3 改进措施	29
4.2 基于多因素影响的非线性时间序列模型	29
4.2.1 气候预测模型	30
4.2.2 多因素作用下的鲁棒变分回声状态网络预测方法	35
4.2.3 算法流程	35
4.3 计算结果	36
4.3.1 验证模型可行性	37
4.3.2 全球气温变化预测	37
5. 问题三：“极寒天气”与“全球变暖”	40
5.1 问题描述	40
5.2 分析思路	40
5.3 有无关系探究：方法 1—地理特征	41
5.4 有无关系探究：方法 2—极值理论	42
5.4.1 相关理论	42
5.4.2 数据来源和处理	44
5.4.3 结果分析	45
5.5 矛盾性探究：分层贝叶斯时空模型	46
5.5.1 相关理论	46
5.5.2 数据来源和处理	48
5.5.3 模型建立	48

5.5.4 模型求解及结果分析.....	50
5.6 本章小结.....	50
6. 问题四：“全球变暖”——沮丧？彷徨？希望？.....	51
6.1 全球气候系统概况.....	51
6.1.1 气候系统的构成.....	51
6.1.2 气候系统内各圈层的相互作用	56
6.1.3 气候系统外的影响因素.....	59
6.1.4 影响气候变化的完整系统.....	59
6.2 从系统动力学角度分析气候变化.....	60
6.2.1 系统动力学理论基础.....	60
6.2.2 气候变化系统的系统分析.....	61
6.2.3 气候变化 SD 模型构建及分析.....	63
6.2.4 如何理解“全球变暖”	64
6.3 气候变化对我们的启示.....	65
7. 参考文献	66
8. 附录	69

1. 问题重述

1.1 问题的背景

2007年2月2日，联合国“政府间气候变化专门委员会”（Intergovernmental Panel on Climate Change, IPCC）发表了第4份全球气候变化评估报告的梗概，这份报告综合了全世界科学家6年来的科学研究成果。报告称气候变暖已经是“毫无争议”的事实。报告指出，从现在开始到2100年，全球平均气温最可能的“升高幅度”是 $1.8\text{--}4^\circ\text{C}$ ，海平面也会因此升高 $18\text{--}59\text{cm}$ 。过去50年全球平均气温上升有超过90%的可能与人类使用化石燃料产生的温室气体增加有关，这是IPCC首次使用这样严重的措辞形容人类活动与气候变暖之间的关联。

2007年4月6日，经过5天的激烈争论，IPCC第二工作组评估报告的决策者摘要终于在比利时的布鲁塞尔面世了。一时间，全球变暖、水资源短缺、洪水、干旱、饥荒、瘟疫暴发、物种灭绝等再次成为流行语。

“温室效应”现象早在1827年就被法国科学家所认识，但“全球气候改变”这个议题是在上世纪初才密集出现。令人遗憾的是，科学家们对“全球气候改变”指的是全球变暖还是变冷抑或暴冷暴热，却莫衷一是。的确，系统的、连续性的大气层尤其是大气对流层温度的数据记载，是在上世纪60年代气象气球探测及微波遥感技术后才出现，距今也不过50年的时间。在这么短的时间跨度内让科学家们准确一致地计算出地球是在“发冷或是发热”，是件让科学家们为难的事，更何况科学的精神历来就崇尚怀疑主义。

因此，一旦出现与“全球变暖”相违背的事件，不仅科学家们纷纷怀疑，人们的直观感受也使得对“全球变暖”这一严重实事放松警惕，比如“冬天太冷”则使人们认为天气并未变暖。然而事实论证了全球变暖实为人类活动引起的观点，并把这一相关性的级别抬高到90%的水平。IPCC委员会指出，在过去的100年中，全球平均地表气温升高 0.74°C ；过去50年的全球平均气温在过去的500年和1300年以来可能是最高的，20世纪的北半球可能是过去1000年中最热的世纪，而90年代又是最温暖的10年，其中1998年和2005年是有记录以来最暖的。另据世界气象组织（WMO）的报告，刚刚过去的2006年全球平均气温较1961-1990年均值高 0.42°C ，为有记录以来的第六暖年。IPCC委员会认为，全球气温在这一时期的异常和快速升高与人类进入温室气体排放密集期正好相吻合，人类活动就是全球变暖的主因。

1.2 要解决的问题

面对“全球变暖”，大部分怀疑者都会有这样的问题：地球真的正在变暖吗？怀疑者声称，我们这个星球没有正在变暖。更确切而言，“目前，也许最不情愿接受的事实是，十多年来全球缺乏变暖”。实际上，气候变暖不是一蹴而就，它是一个大尺度变化的过程，即全球气温在最近一个多世纪以来呈上升趋势，堪称气候科学和统计学史上最确凿的发现之一。

为了进一步验证全球变暖，我们需要解决一下几个问题：

问题一：加拿大各地的温度呈现怎么样的时空变化趋势，这种趋势和全球变暖之间是否存在联系？全球海洋表面温度的历史数据存在怎么样的规律，这种规律与全球变暖之间有无直接或者间接的联系？

问题二：当我们面对密集的数据，不同的分析角度将可能得出不同的结果，面对“全球变暖”的质疑，我们应如何对未来的气候进行预测呢？预测的结果是否会预示着“全球变暖”？

问题三：极端天气的出现是使大众质疑“全球变暖”的最直接原因。极寒天气的出现与全球变暖之间有没有关系？如果有的话，以怎样的形式呈现？并且局地极寒现象与全球变暖的出现之间是否存在矛盾？

问题四：然而解释“全球变暖”与“某地的天气特别冷”的关系，它本身又是一个什么样的概念？是怎么反应气候变化的复杂性及趋势呢？

2. 模型假设与前期任务

为了便于问题的研究，对题目中某些条件进行简化及合理的假设。

针对问题一：

假设各个地区所选气象站的温度为各个地区的平均温度。

针对问题二：

假设（1）：由于气候变化的不确定性，假设每年的气温数据服从概率分布。

假设(2)：由于研究问题的局限性，仅用温度代表气候变化的特征。

假设(3): 由于气候系统的复杂性, 建立简单数学模型时, 仅考虑已知的影响要素。

针对问题四：

假设我们所研究影响气候变化的整个系统存在边界。

为了使研究方向更加精确,本文首先对来自 WOS 的 611 篇有关全球变暖的高被引论文,运用 VOSview 软件进行文献计量分析,得到关键词聚类结果与研究机构分布,分别见图 2-1、图 2-2。

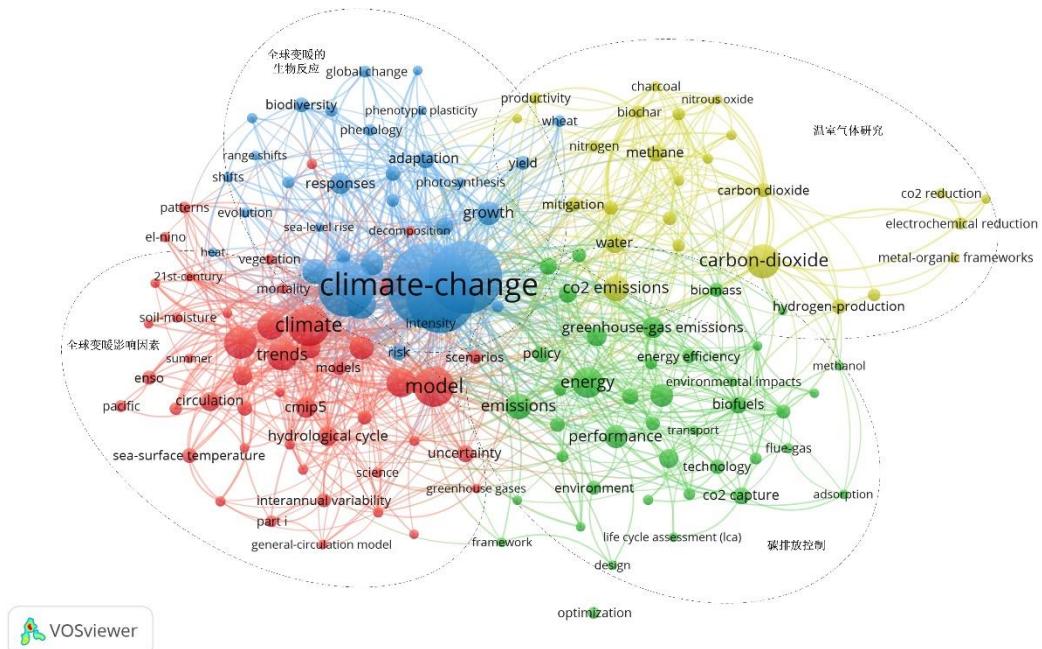


图 2-1 全球变暖研究关键词聚类

由图 2-1 可得，全球变暖高被引论文的研究方向可分成四类，分别是全球变暖的影响因素研究、全球变暖的生物反应研究、温室气体研究以及碳排放控制的研究。

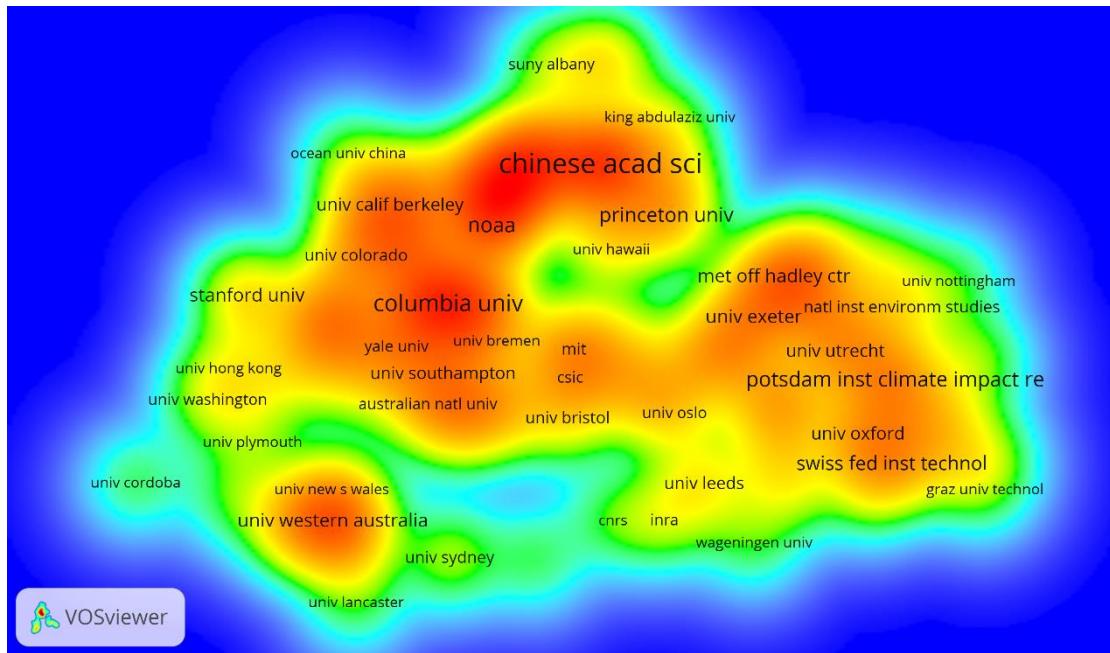


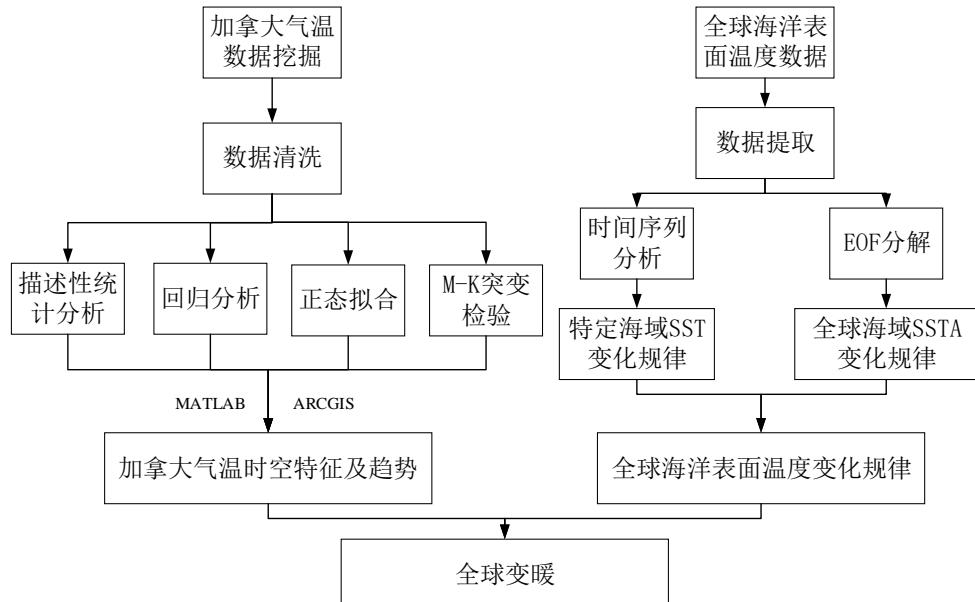
图 2-2 全球变暖研究机构密度图

由图 2-2 可知，全球变暖的研究中心是中国科学院、哥伦比亚大学、西澳大利亚大学以及 NOAA（美国国家海洋大气局）。

3. 问题一：气候数据挖掘

3.1 加拿大各地温度的时空变化趋势

本章具体分析流程如下：



3.1.1 加拿大概况

加拿大位于北美洲北部。东临大西洋，西濒太平洋，西北部邻美国阿拉斯加州，南接

美国本土，北靠北冰洋。海岸线约长 24 万多公里。东部气温稍低，南部气候适中，西部气候温和湿润，北部为寒带苔原气候。中西部最高气温达 40℃以上，北部最低气温低至 -60℃。全国分 10 省 3 地区，10 省为不列颠哥伦比亚、阿尔伯塔、萨斯喀彻温、曼尼托巴、安大略、魁北克、新不伦瑞克、诺瓦斯科舍、爱德华王子岛、纽芬兰和拉布拉多，3 地区为育空、西北、努纳武特。

3.1.2 研究方法

(1) 描述性统计分析

描述性统计，是指运用制表和分类，图形以及计算概括性数据来描述数据特征的各项活动。描述性统计分析要对调查总体所有变量的有关数据进行统计性描述，主要包括数据的频数分析、集中趋势分析、离散程度分析、分布以及一些基本的统计图形。

① 数据的频数分析。在数据的预处理部分，利用频数分析和交叉频数分析可以检验异常值。

② 数据的集中趋势分析。用来反映数据的一般水平，常用的指标有平均值、中位数和众数等。

③ 数据的离散程度分析。主要是用来反映数据之间的差异程度，常用的指标有方差和标准差。

④ 数据的分布。在统计分析中，通常要假设样本所属总体的分布属于正态分布，因此需要用偏度和峰度两个指标来检查样本数据是否符合正态分布。

⑤ 绘制统计图。用图形的形式来表达数据，比用文字表达更清晰、更简明。在 SPSS 软件里，可以很容易地绘制各个变量的统计图形，包括条形图、饼状图和折线图等。

(2) 回归分析

回归分析(Regression Analysis)是确定两种或两种以上变量间相互依赖的定量关系的一种统计分析方法。运用十分广泛，回归分析按照涉及的自变量的多少，分为回归和多重回归分析；按照自变量的多少，可分为一元回归分析和多元回归分析；按照自变量和因变量之间的关系类型，可分为线性回归分析和非线性回归分析。如果在回归分析中，只包括一个自变量和一个因变量，且二者的关系可用一条直线近似表示，这种回归分析称为一元线性回归分析。如果回归分析中包括两个或两个以上的自变量，且因变量和自变量之间是线性关系，则称为多重线性回归分析。用 x_i 表示样本量为 n 的某一气候变量，用 t_i 表示 x_i 所对应的时间，建立 x_i 与 t_i 之间的一元线性回归：

$$x_i = a + bt_i \quad (i=1, 2, \dots, n) \quad (3-1)$$

式中 a 为回归常数， b 为回归系数。 a 和 b 可以用最小二乘法进行估计，对观测数据 x_i 及相应的时间 t_i ，回归系数 b 和常数 a 的最小二乘估计为：

$$\begin{cases} b = \frac{\sum_{i=1}^n x_i t_i - \frac{1}{n} (\sum_{i=1}^n x_i) (\sum_{i=1}^n t_i)}{\sum_{i=1}^n t_i^2 - \frac{1}{n} (\sum_{i=1}^n t_i)^2} \\ a = \bar{x} - b \bar{t} \end{cases} \quad (3-2)$$

其中

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \bar{t} = \frac{1}{n} \sum_{i=1}^n t_i \quad (3-3)$$

回归系数 b 的符号表示气候变量 x 的趋势倾向。 b 值的大小反映了上升或下降的倾向程度。 b 值为正，即当 $b > 0$ 时，说明随着 t 的增高， x 呈上升趋势；当 b 值为负，即当 $b < 0$ 时，说明随着 t 的增高， x 呈下降趋势，因此，通常将 b 称为倾向值，这种方法称为线性倾向估计。

(3) 正态拟合

正态分布概念是德国数学家 Moivre 在 1733 年首次提出，著名德国数学家高斯率先推导出正态分布概率密度公式。正态分布是自然科学与社会科学中都非常重要的概率分布，各种各样的心理学、经济学等社会科学测试以及数学、物理等自然科学测试现象都被发现近似地服从正态分布。正态分布是在统计学中应用最广泛的一种概率分布。正态分布的概率分布函数见公式 3-3

$$f(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (3-3)$$

(4) Mann-Kendall 检验

Mann-Kendall 的检验方法是非参数方法。非参数检验方法亦称为无分布检验，其优点是不需要样本遵从一定的分布，也不受少数异常值的干扰，更适用于类型变量和顺序变量，计算也比较简便。具体方法如下。

对于具有 n 个样本量的时间序列 X ，构造秩序列：

$$s_k = \sum_{i=1}^k r_i \quad r_i = \begin{cases} 1 & x_i > x_j \\ 0 & \text{else} \end{cases} \quad j = 1, 2, \dots, i \quad (3-4)$$

可见，秩序列 s_k 是第 i 时刻数值大于 j 时刻数值个数的累计数。在时间序列随即独立的假定下，定义统计量

$$UF_k = \frac{s_k - E(s_k)}{\sqrt{Var(s_k)}} \quad k = 1, 2, \dots, n \quad (3-5)$$

式中 $UF_1=0$, $E(s_k)$, $Var(s_k)$ 是累计数 s_k 的均值和方差，在 x_1, x_2, \dots, x_n 相互独立且有相同连续分布时，它们可由下式给出：

$$E(s_k) = \frac{n(n+1)}{4} \quad Var(s_k) = \frac{n(n-1)(2n+5)}{72} \quad (3-6)$$

UF_i 为标准正态分布，它是按时间序列 x 的顺序 x_1, x_2, \dots, x_n 计算出来的统计量序列，给定显著性水平 α ，查正态分布表，若 $|UF_i| > U_\alpha$ ，则表明序列存在明显的趋势变化。

按时间序列 x 的逆序 x_n, x_{n-1}, \dots, x_1 ，再重复上述过程，同时使

$$UB_k = -UF_k, k = n, n-1, \dots, 1 \quad UB_1 = 0 \quad (3-7)$$

3.1.3 数据处理

本题所采用的气象数据来自于附件中所给的加拿大政府网站。为了简化分析过程，在加拿大 10 省 3 地区中各自选择一个具有代表性的气象站点，获取该站点所在省或者地区

的 2000-2019 年逐日气象数据。选取的气象站台分别是 ATHABASCA 2(阿尔伯塔)、HOWE SOUND - PAM ROCKS (不列颠哥伦比亚)、ROBLIN (曼尼托巴)、MIRAMICHI A (新不伦瑞克)、GANDER INTL A (纽芬兰和拉布拉多)、LAKE MAJOR (诺瓦斯科舍)、FORT LIARD (西北)、DEWAR LAKES (努纳武特)、BROCKVILLE PCC (安大略)、CHARLOTTETOWN A (爱德华王子岛)、LA BAIE (魁北克)、ROCK POINT (萨斯喀彻温)、ROCK RIVER (育空)，由于这 13 个气象站点的气象数据再研究时间区间内比较完整，并且均匀的分布在整个研究区域内，可以较好的代表研究地区整体气温及变化趋势各气象站点空间分布见图 3-1，图中颜色仅用于区别各地区。在结果分析中，加拿大各地区由各自的英文缩写表示，具体见表 3-1。

表 3-1 加拿大各地区缩写

地区	缩写	地区	缩写
阿尔伯塔省	AB	努纳武特地区	NU
不列颠哥伦比亚省	BC	安大略省	ON
曼尼托巴省	MB	爱德华王子岛省	PE
新不伦瑞克省	NB	魁北克省	QC
纽芬兰与拉布拉多省	NL	萨斯喀彻温省	SK
新斯科舍省	NS	育空地区	YT
西北地区	NT		

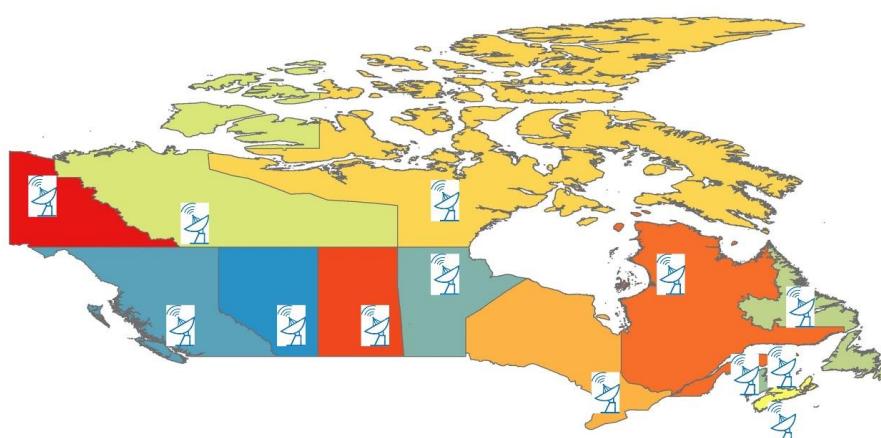


图 3-1 各气象站点分布

然后运用 SPSS 软件对所缺失的气温资料进行数值插补：如果缺少独立一日的数值，则采用缺失值所在的前一日和后一日的平均值来代替；如果缺少一系列的数值，则采用多元回归的方法，选取与缺失值所在的站点相关性最好的五个站点进行插补；然后将补充完整的数值进行计算，得到年、月、四季和年代际的气温、最低气温、最高气温。

3.1.4 分析结果

(1) 加拿大各地区 2000-2019 年平均气温描述统计分析

运用 SPSS 软件对加拿大 13 个地区 2000-2019 年平均气温计算平均值、最大值、最小值以及标准偏差等，结果见表 3-2。全时间段平均气温的空间分布图见图 3-2。

表 3-2 加拿大各地区 2000-2019 年平均气温描述统计分析结果

地区	最小值	最大值	平均值	标准偏差
AB	1.25	5.1	2.87	1.00
BC	10.27	12.77	11.36	0.62
MB	-0.57	3	1.41	1.07
NB	4.52	7.23	5.58	0.77
NL	3.76	6.14	4.82	0.59
NS	4.94	8.84	7.42	0.92
NT	-1.64	1.5	-0.36	0.84
NU	-15.43	-7.54	-11.78	1.65
ON	6.31	9.44	8.00	0.87
PE	5.35	7.6	6.24	0.63
QC	1.23	4.48	2.60	0.86
SK	2.56	5.56	3.99	0.89
YT	-6.5	-0.11	-3.86	1.69

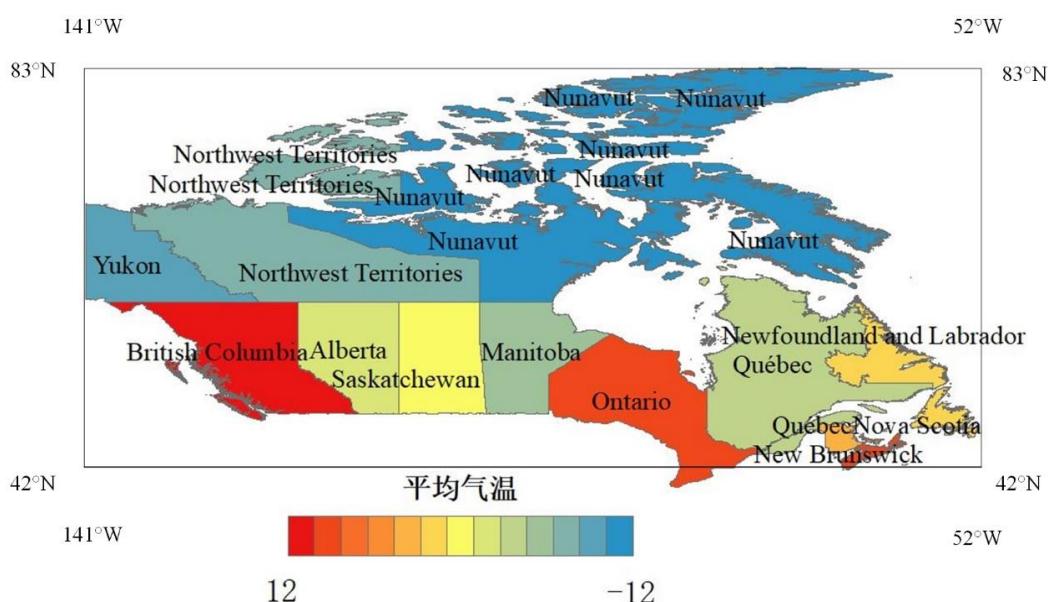


图 3-2 加拿大各地区平均气温分布

由表 3-2 与图 3-2 可得，加拿大年全时间段平均气温最低的是努纳武特地区，位于加拿大的东北地区，该地区全境多为冰雪所覆盖，DEWAR LAKES 统计站 2000-2019 年统计该地区共出现 10 次“极寒天气”，最低日平均气温为 2004 年 3 月 10 日出现的 -42.9°C 。而平均气温最高的地区是不列颠哥伦比亚省，位于加拿大西南地区，邻近太平洋。北部三个地区的全时间段平均温度均低于 0°C 。同时可以看出同样纬度，邻近太平洋或者大西洋的地区，其平均气温略高于大部分其他地区，如不列颠哥伦比亚省与新斯科舍省。

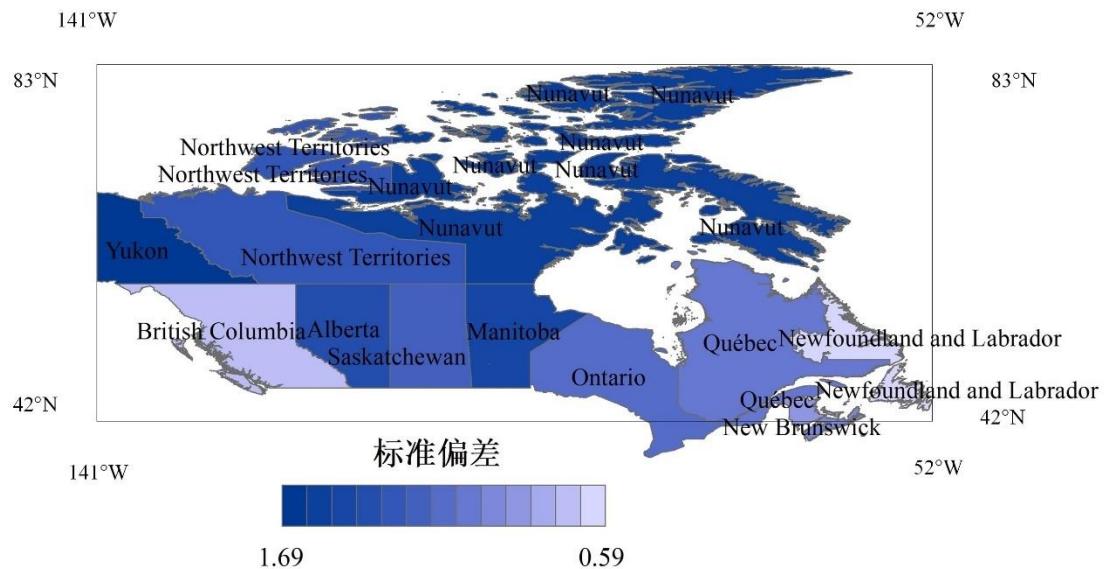
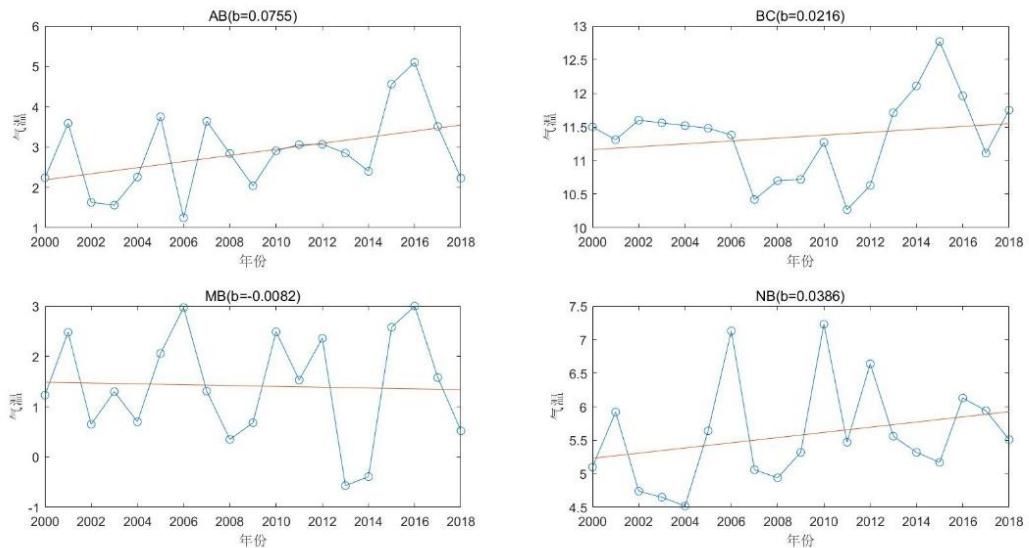


图 3-3 加拿大各地区年平均气温标准偏差

年平均气温的标准偏差可以说明各个地区年平均气温的变化程度，由图 3-3 可得，不列颠哥伦比亚省的标准偏差只有 0.59，即该地区的年平均气温变化范围较小，基本保持在均值上下轻微浮动，其 2000-2019 年平均气温的最大值与最小值只相差 2.5°C ，而标准偏差最大的省份为育空地区与努纳武特地区，其 2000-2019 年平均气温的最大值与最小值相差 6.39°C 与 7.89°C 。

(2) 加拿大各地区 2000-2019 年平均气温时间序列分析

将加拿大 13 个地区以及加拿大整个国家的 2000-2019 年平均气温以时间序列方式展开，如图 3-4，图中 b 表示回归系数。



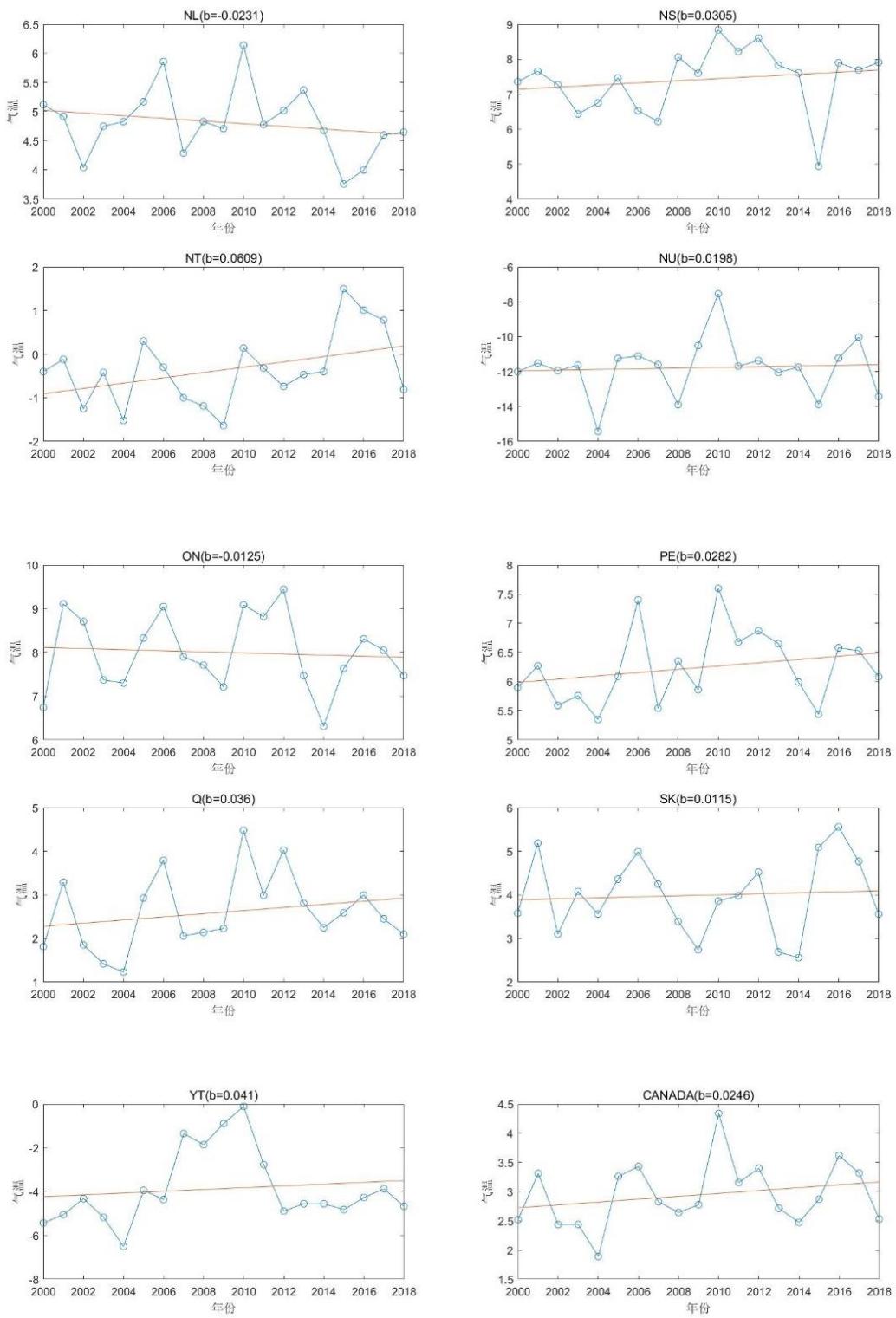
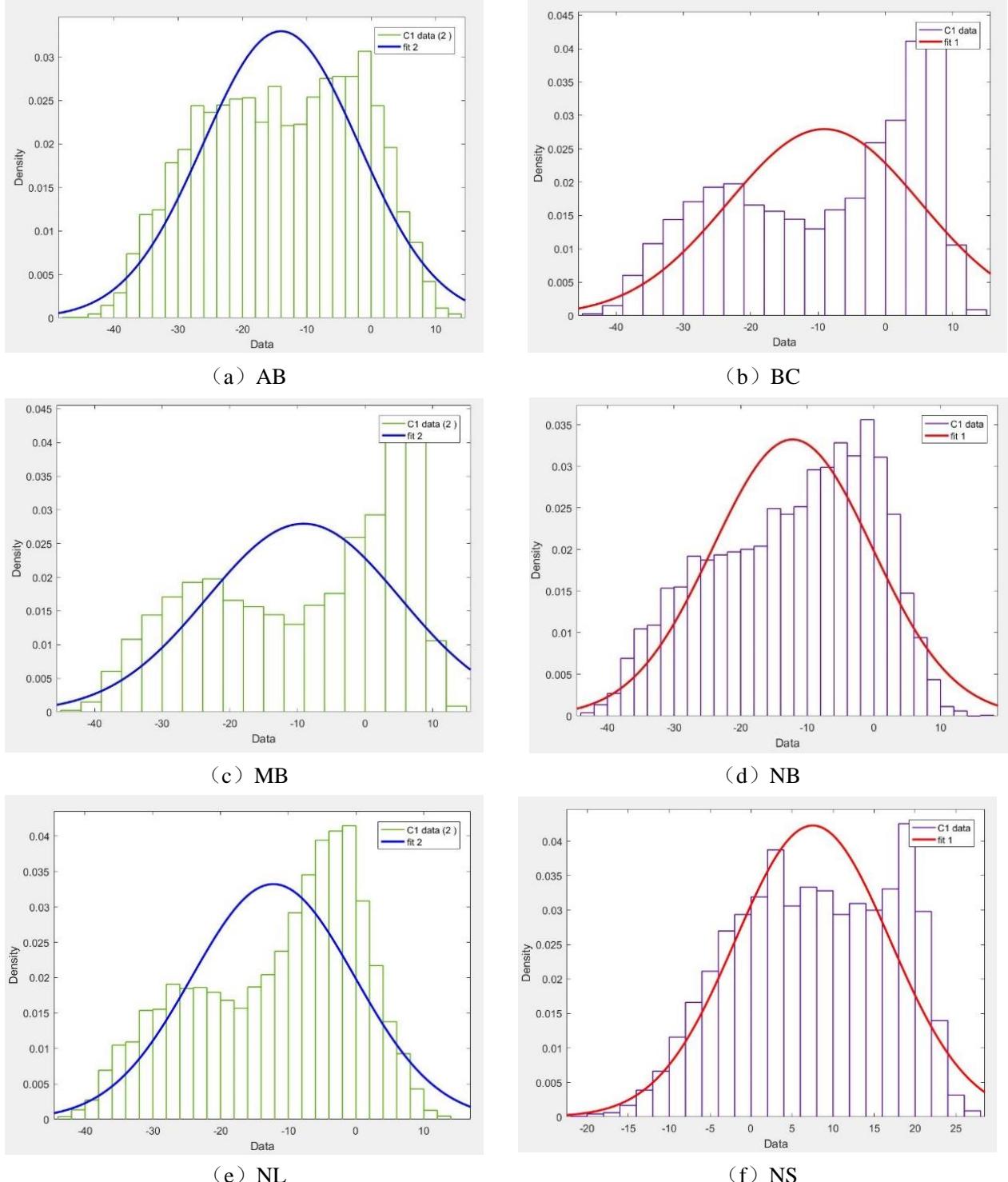


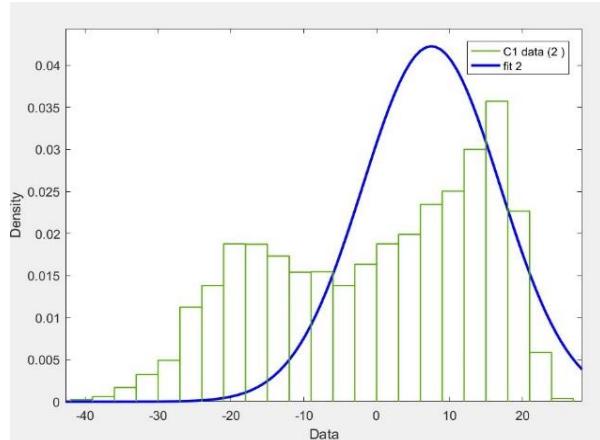
图 3-4 加拿大及其 13 个地区 2000-2019 年平均气温

由图 3-4 可得，加拿大 13 个地区时间序列变化曲线的回归系数中，有 10 个地区大于 0，即这十个地区的平均气温在 2000-2019 年间呈上升趋势，但上升幅度较小，MB、NL、ON 地区的平均气温在 2000-2019 年间呈下降趋势，下降幅度也较小。加拿大全国的平均气温在 2000-2019 年间呈上升趋势。

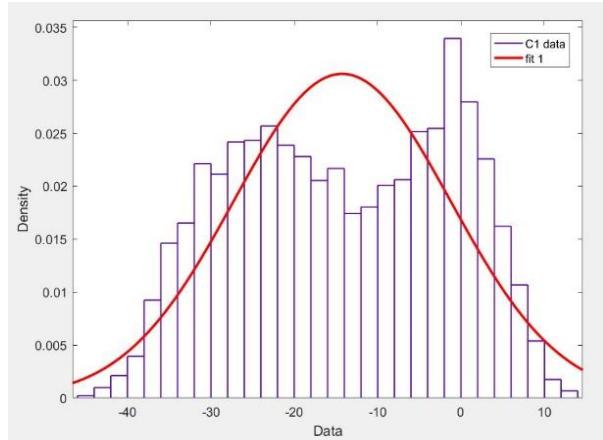
(3) 加拿大各地 2000-2019 年日气温正态拟合

运用 MATLAB 的 Distribution Fitting 工具箱对加拿大 13 个地区的日气温分布进行正态拟合，拟合结果见图 3-5。由图可得，这些气温观测数据中，AB、BC、MB、NB、NL、NU 这些地区均出现过低于 -40°C 的“极寒天气”，所有地区的气温分布均较符合正态分布特征，以平均气温为均值，极少数情况出现高于平均气温太多或者低于平均太多的极端天气。

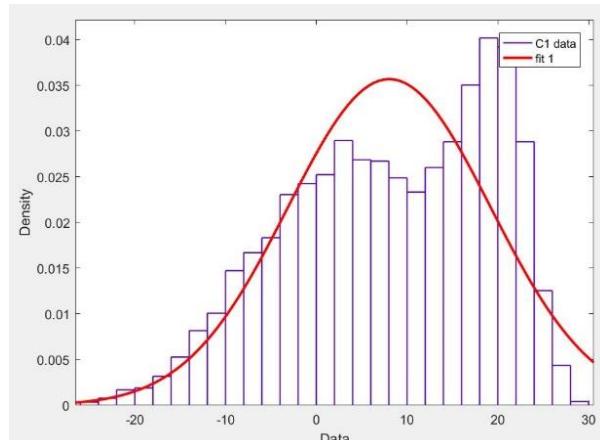




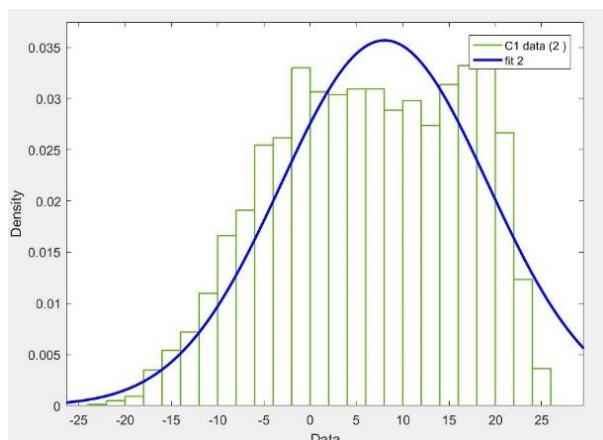
(g) NT



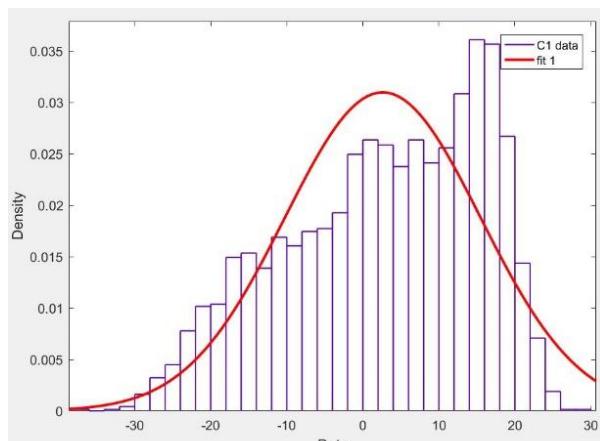
(h) NU



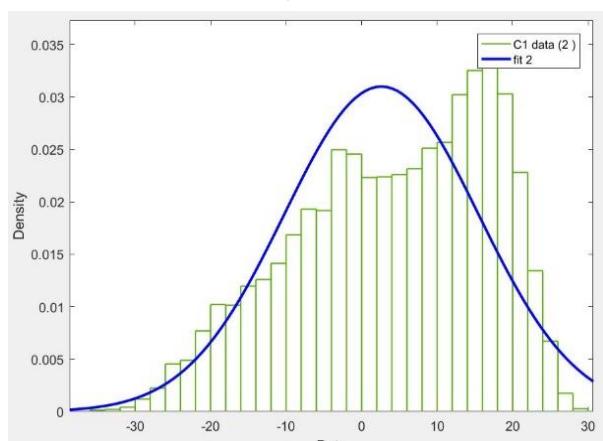
(i) ON



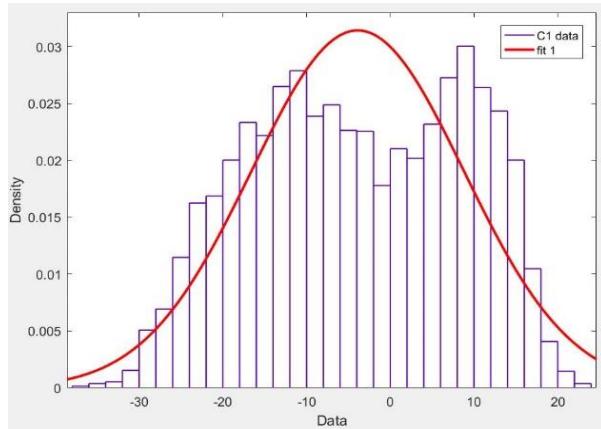
(j) PE



(k) QC



(l) SK



(m) YT

图 3-5 加拿大各地区日平均气温正态拟合

(4) 加拿大各地区 2000-2019 年夏季、冬季平均气温变化趋势

根据文献调查法，本节中的加拿大各地区的夏季平均气温统一用 6、7、8 月的平均气温代替，冬季平均气温用 1、2、3 月份的平均气温代替。其中夏季平均气温的地理分布如图 3-6。夏季平均气温最高的地区是安大略省，平均气温为 19.7°C 。

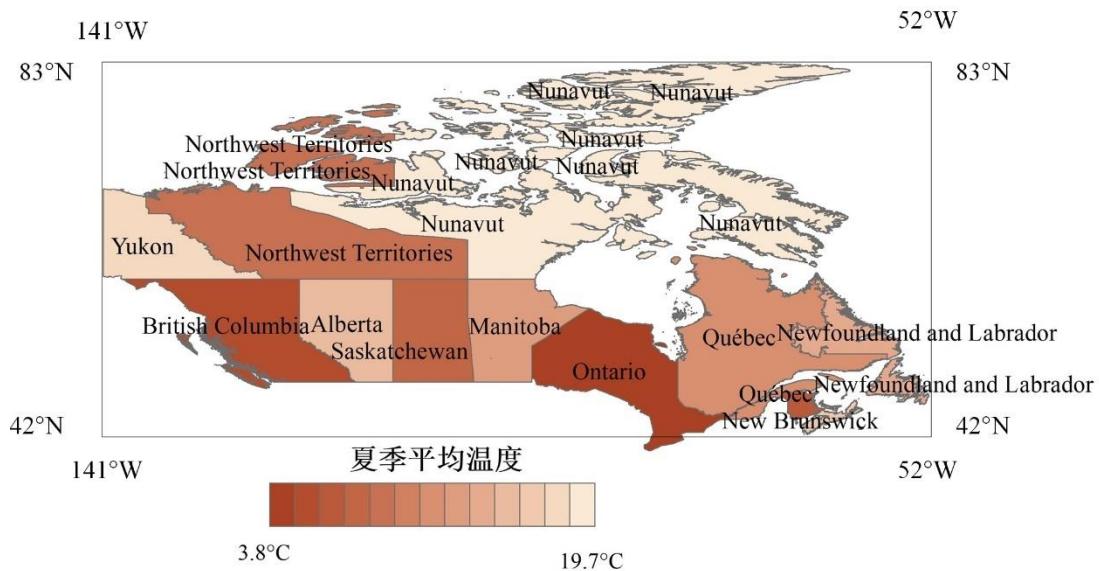


图 3-6 加拿大各地区夏季平均温度

图 3-7 为加拿大各地区冬季平均温度，如图中所示，颜色越浅表示冬季平均温度约高，颜色越深表明冬季平均温度约低，其中努纳武特地区的冬季平均气温低至 -26°C

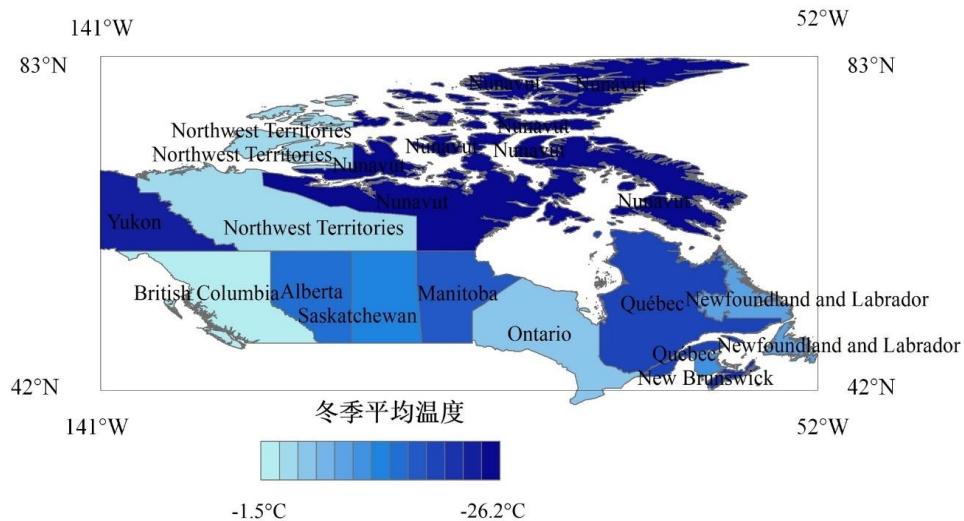


图 3-7 加拿大各地区冬季平均温度

(5) 加拿大 2000-2018 年平均气温 MK 突变检验分析

利用 Mann-Kendall 对黑龙江省年平均气温作突变分析, 图 3-8 为黑龙江省年平均气温突变分析结果。如图所示, UF 与 UB 在置信区间内有 2 个明显的交点, 说明在这两年发生突变, 根据 UF 与 UB 交点的位置, 确定这两年是 2001 年和 2005 年。在 2001 年 -2005 年期间 $UF < 0$, 说明气温呈现出下降的趋势; 在 2005 年以后, UF 始终大于 0, 气温升高且达到最高值, 增温趋势变得明显。

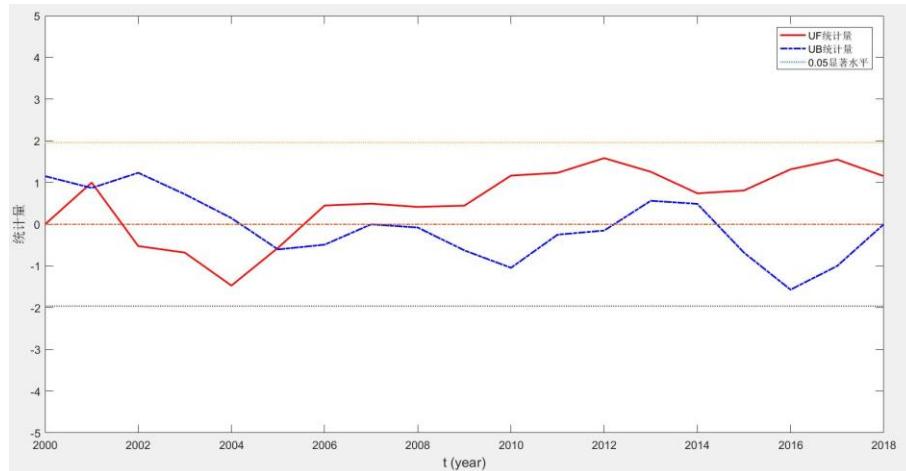


图 3-8 加拿大 2000-2018 年平均气温 Mann-Kendall 突变检验分析图

3.2 海洋表面温度数据规律

3.2.1 研究背景

海表温度 (Sea Surface Temperature, SST) 是海洋与大气之间相互作用的基础, 海-气之间能量的交换正是通过海-气界面进行的, SST 的任何微小变化都可能标志着海洋内部热能储蓄的重大变化, SST 表征了海洋热力、动力过程, 是海洋与大气相互作用的综合结果, 是反映海气热量、动力和水汽交换的重要参量, 是气候变化的主要影响因素和参考指标, 对厄尔尼诺-南方涛动 (El Niño Southern Oscillation, ENSO) 的发生、发展扮演着重要角色。观测与数值模拟表明, 经度 $20^\circ \times$ 纬度 20° 以上大尺度的海表面温度异常能引起大气环流的显著变化, 如厄尔尼诺和拉尼娜现象。在厄尔尼诺期间, 热带东太平洋信风减弱, 该海域海面水温大幅度上升, 比正常年份升高 $3\text{--}5^\circ\text{C}$, 从而使得大气环流和大洋环发

生重大改变，造成世界范围内的大气和海洋生态环境以及气候的异常。

3.2.2 相关理论

(1) ENSO

当厄尔尼诺现象发生时，赤道东太平洋大范围的海水温度可比常年高出几摄氏度。太平洋广大水域的水温升高，改变了传统的赤道洋流和东南信风，使全球大气环流模式发生变化，其中最直接的现象是赤道西太平洋与印度洋之间海平面气压的成反相关，即南方涛动现象（Southern Oscillation, SO）。在拉尼娜现象期间，东南太平洋气压明显升高，印度尼西亚和澳大利亚的气压减弱。厄尔尼诺期间的情况正好相反。这种海洋与大气的相互作用现象，气象上把两者合称为 ENSO(El Niño/La Niña-Southern Oscillation，简写 ENSO，音“恩索”）。这种全球尺度的气候振荡被称为 ENSO 循环。

(2) 拉尼娜现象

拉尼娜现象就是太平洋中东部海水异常变冷的情况。东南信风将表面被太阳晒热的海水吹向太平洋西部，致使西部比东部海平面增高将近 60 厘米，西部海水温度增高，气压下降，潮湿空气积累形成台风和热带风暴，东部底层海水上翻，致使东太平洋海水变冷。太平洋上空的大气环流叫做沃克环流，当沃克环流变弱时，海水吹不到西部，太平洋东部海水变暖，就是厄尔尼诺现象；但当沃克环流变得异常强烈，就产生拉尼娜现象。一般拉尼娜现象会随着厄尔尼诺现象而来，出现厄尔尼诺现象的第二年，都会出现拉尼娜现象，有时拉尼娜现象会持续两、三年。1988-1989 年，1998-2001 年都发生了强烈的拉尼娜现象，1995-1996 年发生的拉尼娜现象较弱，有的科学家认为，由于全球变暖的趋势，拉尼娜现象有减弱的趋势。2011 年拉尼娜现象在赤道太平洋海域开始加强。拉尼娜与厄尔尼诺性格相反，随着厄尔尼诺的消失，拉尼娜的到来，全球许多地区的天气与气候灾害也将发生转变。总体说来，拉尼娜并非性情十分温和，它也将可能给全球许多地区带来灾害，其气候影响与厄尔尼诺大致相反，但其强度和影响程度不如厄尔尼诺。

3.2.3 研究方法

(1) SST 时间序列

在全球海域选择一个剖面，对该剖面的海洋表面温度进行时间序列分析。本节选取加拿大西部沿海的 $134^{\circ}\text{W}, 46^{\circ}\text{N}$ 范围内的海洋表面温度作温度一时间序列剖面分析。

(2) 经验正交函数分析法 (EOF)

对于 SST 时间序列的变化，大多数的研究中使用的是主成分分析方法，例 EOF 分解和奇异值分解，或二者的变种。本文使用 EOF 方法对 SST 场进行时空特征分析。EOF 方法，也称特征向量分析（Eigenvector Analysis），属于主成分分析方法的一种，于二十世纪中叶开始被应用于气候和气象研究，现在普遍应用于地学等研究领域中。通过 EOF 分解能够提取数据主要特征量，并对矩阵数据中的结构特征进行分析。EOF 的分解结果中包括不随空间变化的时间函数，即表征时间变化的主成分，也称时间系数；不随时间变化的空间函数，即表征空间分布的特征向量，也称空间模态或空间特征向量，因此地学中也将 EOF 分析称为时空分解。使用这种方法可以达到降维分析的目的，所以对气象学数据的分析具有优势。

由月 SSTA 场和 $2^{\circ} \times 2^{\circ}$ 网格组成原始数据集，为计算方便，将三维的 SST 数据降维成 16020×1985 的数据集，再剔除其中没有数据的网格（陆地），得到矩阵 W_{10988} ，其中行表示每个空间数据点，列表示各数据点每月的温度变化，然后逐行求距平，得到距平矩阵。

计算 W 与其转置矩阵 W^T 的交叉积，得到协方差阵

$$M_{10988 \times 10988} = WW^T / 1985 \quad (3-8)$$

计算 M 的特征根 ($\lambda_1, \dots, \lambda_{10988}$) 和特征向量 $L_{10988 \times 10988}$ 二者满足

$$M_{10988 \times 10988} \times L_{10988 \times 10988} = L_{10988 \times 10988} \times K_{10988 \times 10988} \quad (3-9)$$

$$K = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_m \end{bmatrix} \quad (3-10)$$

特征根 λ 满足 $\lambda_1 > \lambda_2 > \dots > \lambda_{10988}$ 。每列特征向量与一个非 0 的特征根对应，称该特征向量为 EOF，将 EOF 投影到原始数据 W 上，计算出时间系数（即主成分），时间系数与空间特征向量相对应，即

$$XM_{10988 \times 1985} = L_{10988 \times 10988}^T \times W_{10988 \times 1985} \quad (3-11)$$

XM 的每一行数值就是每个特征向量对应的时间系数。

计算主成分的各个模态的贡献率，可以用 λ 的大小来表示矩阵 W 的方差大小， λ 越大表示其对应的模态总方差的贡献越大。第 k 个模态对总的方差解释率为

$$\frac{\lambda_k}{\sum_{i=1}^m \lambda_i} \times 100\% \quad (3-12)$$

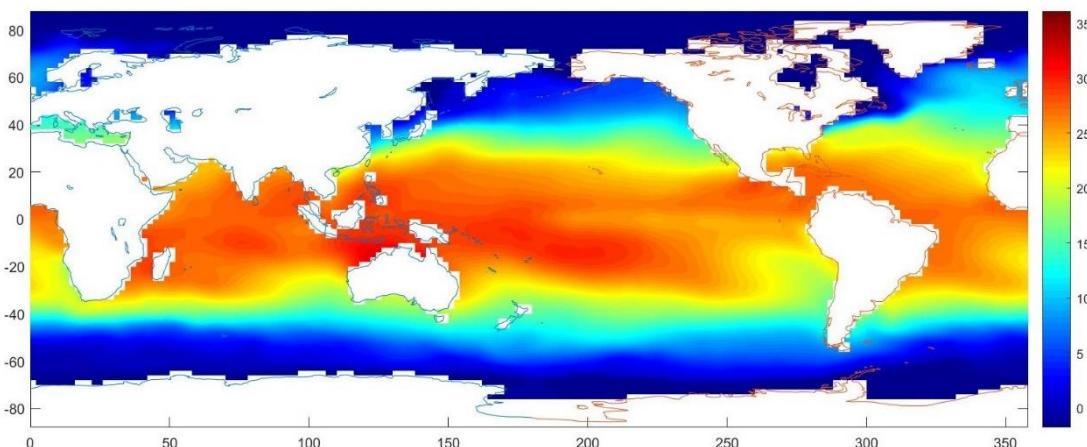
3.2.4 数据处理

本节采用的 SST 观测数据库为美国国家海洋大气管理局的扩建海温 ERSST.v3b (<https://www.esrl.noaa.gov/psd/>) 的月平均数据，空间分辨率为 $2^\circ \times 2^\circ$ ，所用时间范围 1854/01-2019/05。研究的范围是全球海域。

3.2.5 分析结果

(1) 全球海域 SST 特征

选择 1984 年 1 月、1984 年 7 月，以及 2018 年 1 月、2018 年 7 月的全球海域 SST 数据进行分析。得到的月平均 SST 见图 3-9。



(a) 1954 年 1 月

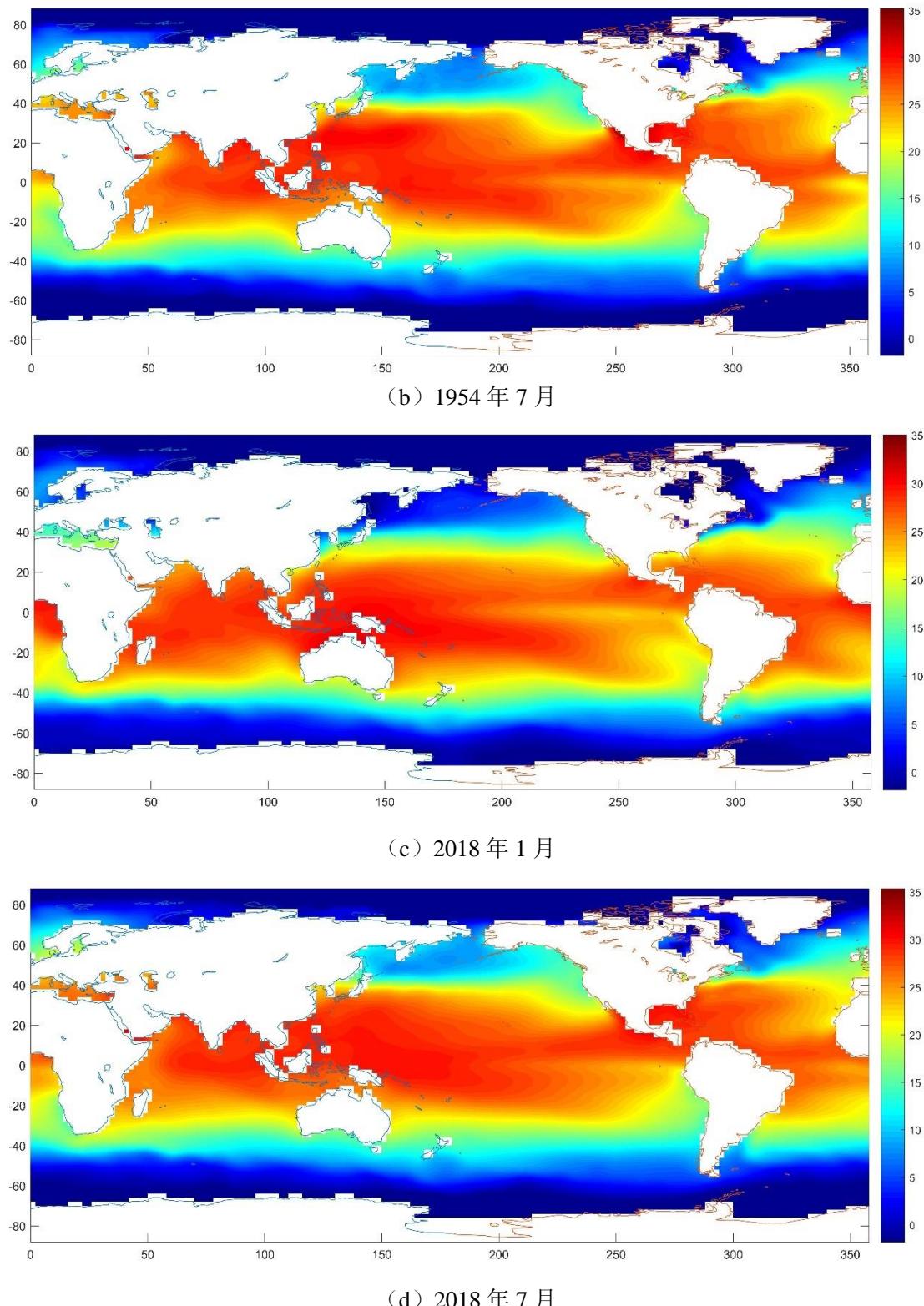


图 3-9 全球海洋表面温度变化

(2) 特定海域时间序列

通过读取研究海域 SST 在 1985 个月的变化，得到时间序列图如图 3-10。图中可见该区域海温的季节性变化多在 20°C 与 28°C 之间，较强的海表面温度信号多集中在夏半年。

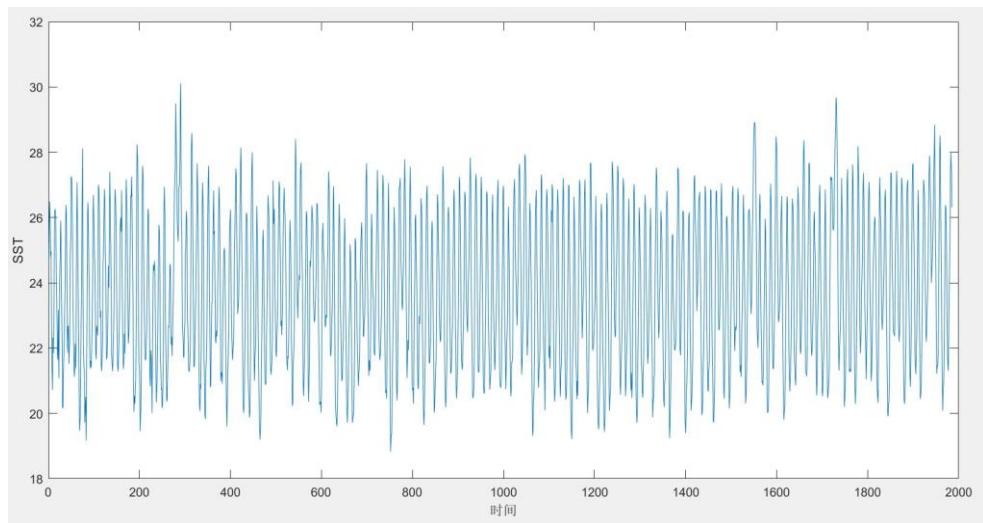


图 3-10 134°W,46°N 海域 SST 变化时间序列

该序列存在季节因素，因此为了更好地说明海洋表面温度的变化，去除季节性影响，按年份进行时间序列分析见图 3-11。由图可知 $b > 0$ ，即该海域 SST 在 1854-2018 年间呈上升趋势。

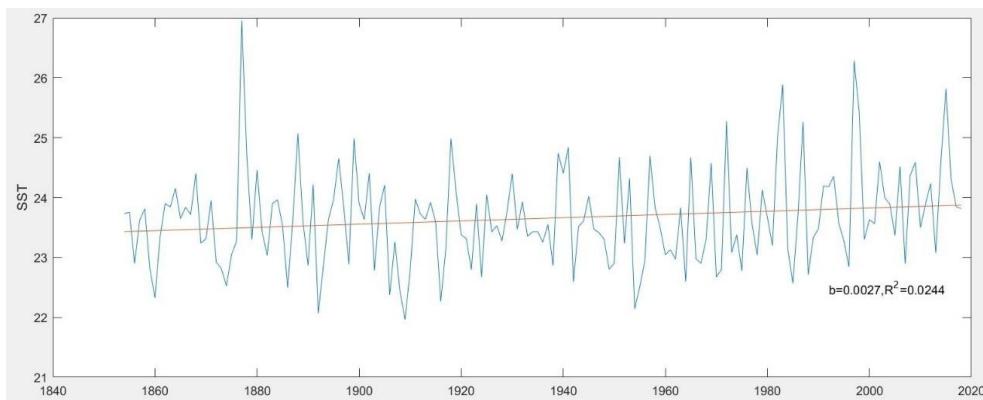


图 3-11 134°W,46°N 海域 SST 年变化时间序列

(3) EOF 分析

将月 SSTA (SST 异常场) EOF 分解得到的时间向量和空间序列按特征值大小排列，表 3-3 为前 5 个模态的方差贡献，其中前五个模态的累积方差贡献率达到 90.99%，因此这 5 个模态基本可以反映全球 SST 的变化。

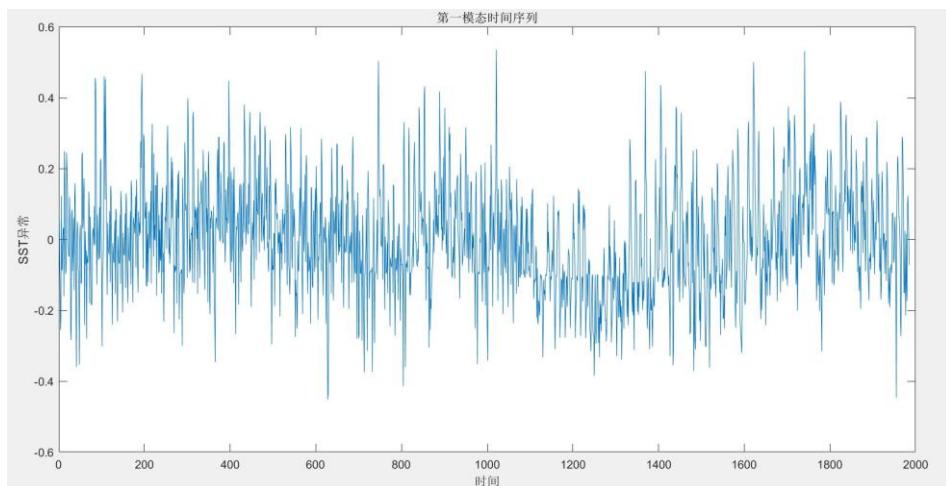
表 3-3 全球 SST 前 5 个模态的方差贡献率

模态	1	2	3	4	5
方差贡献 (%)	82.91	3.16	2.75	1.34	0.83
累积方差贡献 (%)	82.91	86.07	88.82	90.16	90.99

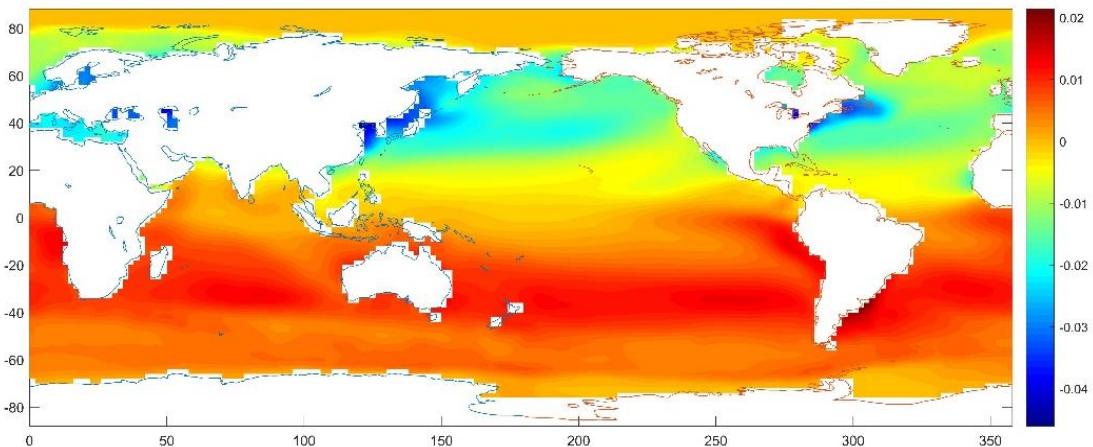
① 第一模态

第一模态在时间上呈季节性波动且呈现出一定的周期性，整体趋势保持平稳。在空间上，南半球大部分地区呈现正异常，而在太平洋西部、大西洋西部存在小部分的负异常。具体表现为，南半球出现异常热的天气，而北半球的太平洋西部与大西洋西部可能出现极寒天气。根据相关文献，这可能是由于北极冰盖融化，最新模型研究表明，如果格林兰冰盖融化可导致其附近海平面将下降 100m，北苏格兰的海平面将下降 3m，冰岛周围海平面将下降 10m，南美部分地区海平面将上升 10m。在海平面附近，大气等位面的变化幅度与

海平面变化幅度非常接近。引力模型的计算结果表明：北极冰盖大量融化导致北极地区海平面和大气等位面的大幅度下降，压力变化迫使北极地区冷水和冷空气流向北半球中低纬度地区。这是美国极寒发生的原因。与此同时，南半球海平面反向上升，赤道热水南移，从低纬度直到南极大陆。这是澳大利亚创纪录高温和南极大陆海冰为南极海冰有记录以来的第二小的1月份的原因。南极大陆海冰在2019年1月份创第二小纪录，南极冰盖融化同样会使地表等位面北移，导致南半球中低纬度变冷，北半球变暖。北极变暖海冰融化，形成又一个冷热循环。这一循环最突出的特征就是，南热北冷和南冷北热循环发生，地表等位面南北震荡周期循环。有资料表明，2019年1月发生美国极寒，2月发生欧洲极热，3月又发生加拿大极寒。地表等位面南北震荡是最合理的解释。



(a) 第一模态时间序列

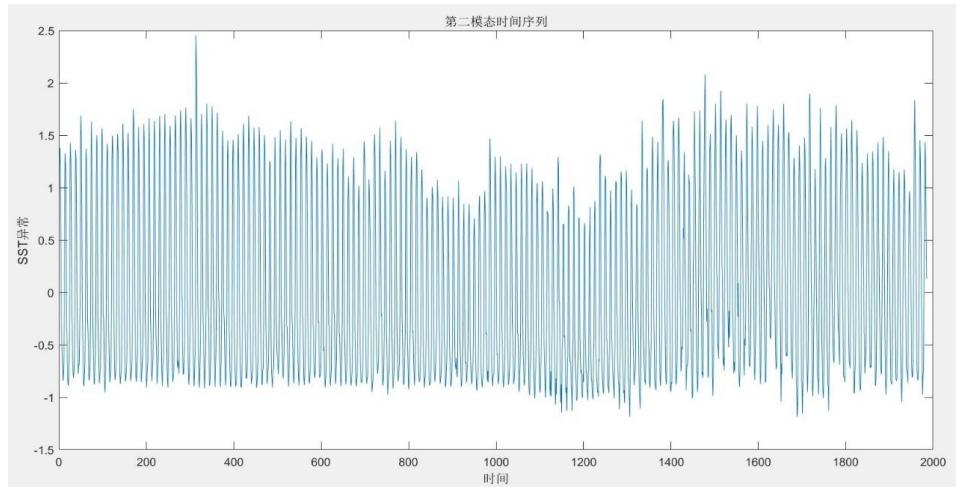


(b) 第一模态空间分布

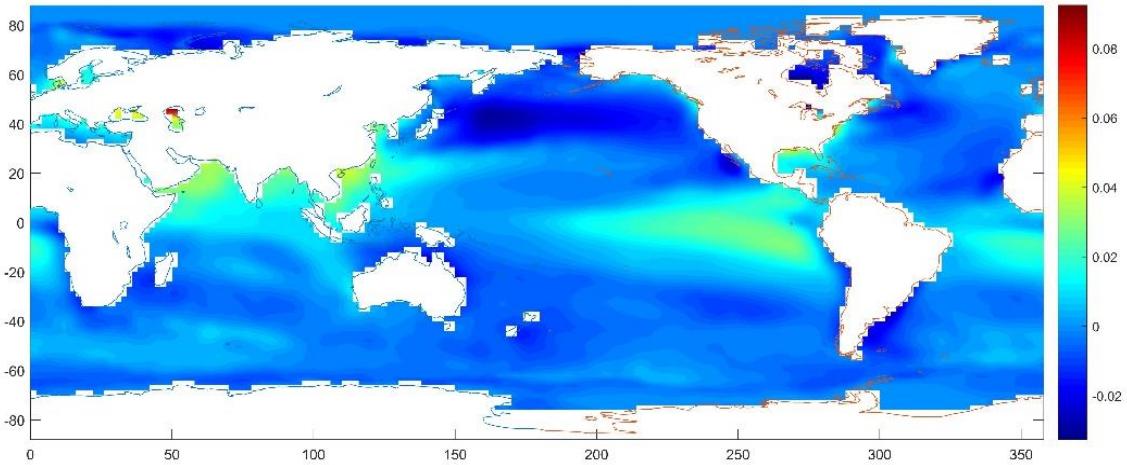
图 3-12 全球 SST 第一模态 EOF 分解时空分布

② 第二模态

第二模态的时间序列呈现明显的周期性。第二模态的空间分布显示，赤道东太平洋、印度洋以及美国东海岸出现正异常，而西太平洋出现负异常，这是典型的 ENSO 模态，即厄尔尼诺现象，赤道中东太平洋是正异常的高值中心，印度洋和大西洋同样表现为正异常，但不如太平洋显著；赤道西太平洋正异常较弱或表现为负异常。



(a) 第二模态时间序列

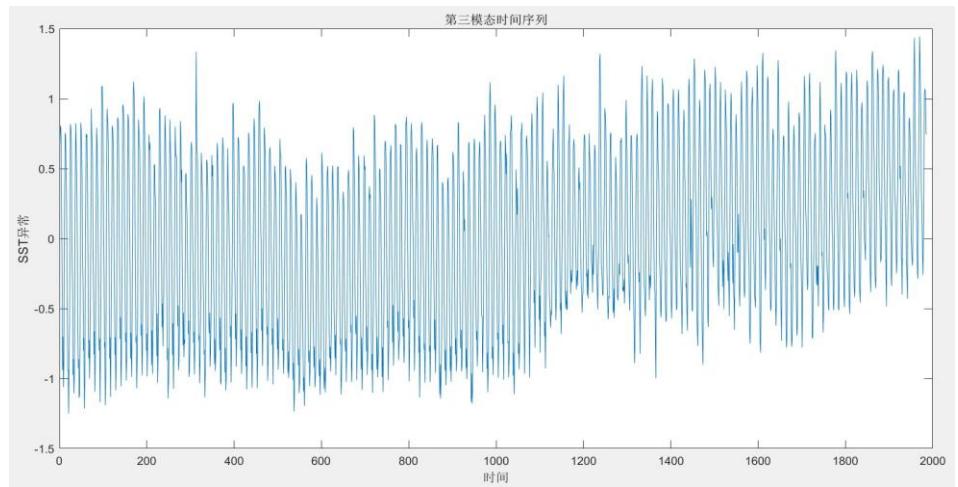


(b) 第二模态空间分布

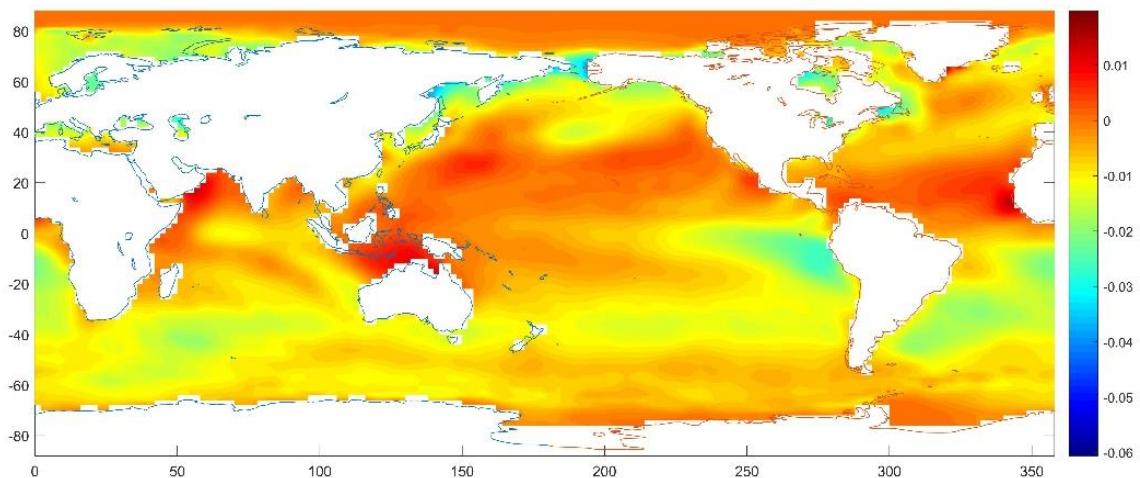
图 3-13 全球 SST 第二模态 EOF 分解时空分布

③ 第三模态

第三模态的时间序列呈现明显的上升趋势，空间分布表明，全球大部分海域的 SST 呈现正异常，这是全球变暖最直接的表现，而时间序列的上升趋势说明这种全球性的海平面温度升高正在加剧。



(a) 第三模态时间序列

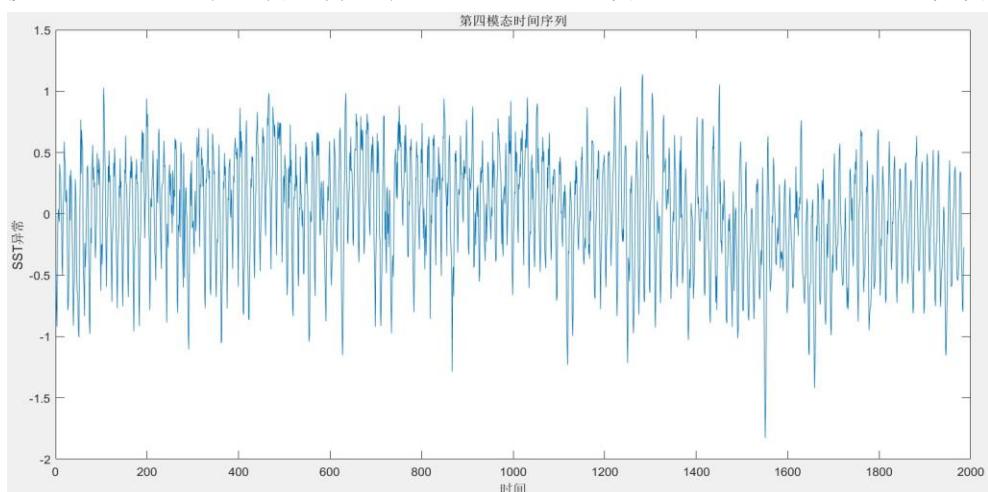


(b) 第三模态空间分布

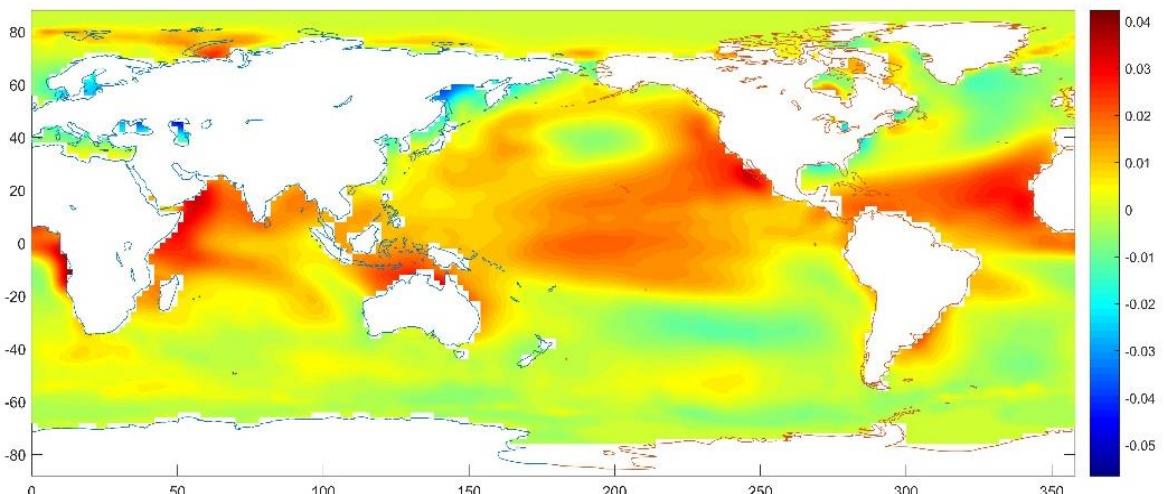
图 3-14 全球 SST 第二模态 EOF 分解时空分布

④ 第四模态

第四模态的空间分布表明，赤道附近地区呈现正异常，西北太平洋出现负异常。



(a) 第四模态时间序列

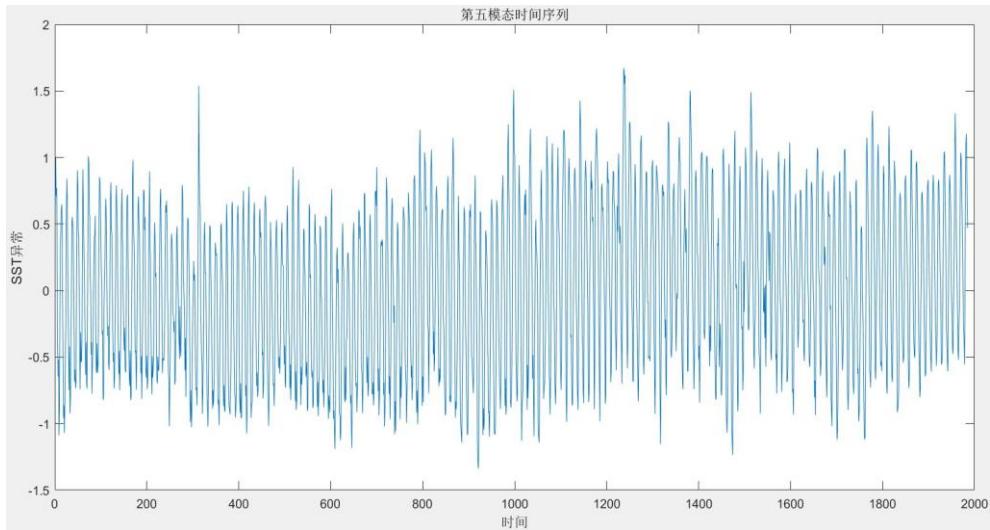


(b) 第四模态空间分布

图 3-15 全球 SST 第四模态 EOF 分解时空分布

⑤ 第五模态

第五模态表现为赤道东太平洋出现负异常，赤道东太平洋出现正异常，这是典型的拉尼娜现象的模态。该模态在时间序列上呈现上升趋势。



(a) 第五模态时间序列

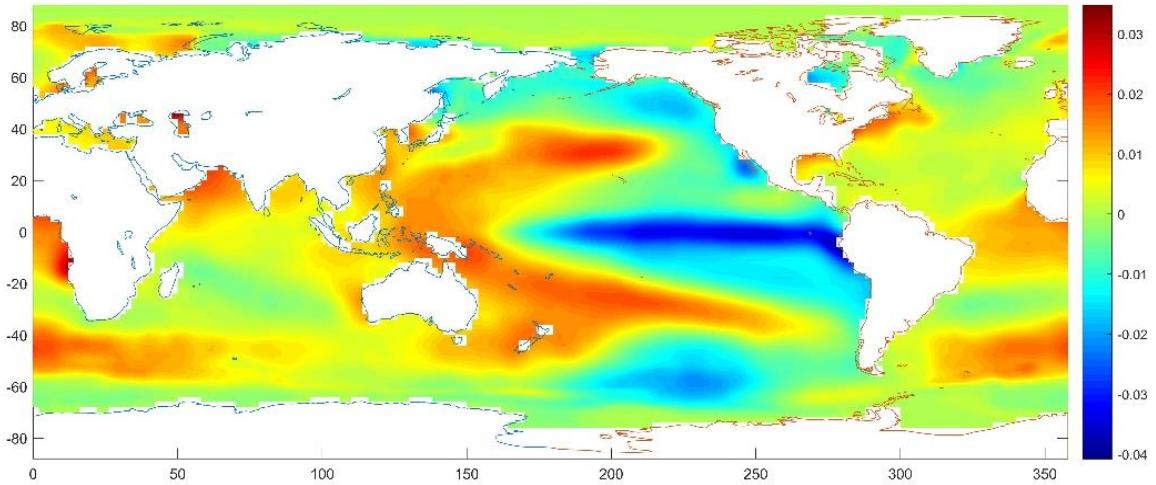


图 3-16 全球 SST 第四模态 EOF 分解时空分布

3.3 本章小结

本章首先从加拿大政府网站挖掘出 13 个地区的气温数据，经过数据清洗、整理，通过描述性统计分析方法、空间分析方法、回归分析方法、正态拟合方法发掘出加拿大地区的气候变化特征。结果显示，加拿大平均气温最高的地区是不列颠哥伦比亚省，平均气温最低的是努纳武特地区，年平均气温变化最大的是育空地区与努纳武特地区。加拿大大部分地区以及加拿大的年平均气温有上升趋势，AB、BC、MB、NB、NL、NU 这些地区均出现过低于 -40°C 的“极寒天气”。2001 年和 2005 年的平均气温是两个突变值，且在 2001-2005 年加拿大年平均气温有下降趋势，但 2005 年之后一直保持上升趋势。

对全球 SST 的数据进行时间序列分析以及 EOF 分析，结果显示 1854 年的全球海洋表面温度分布与 2018 年的全球海洋表面温度分布无明显差别。加拿大西南海域的年平均气温保持在 22°C 与 27°C 之间。EOF 分解的第一模态表现为南半球海域异常升温，西太平洋与北大西洋海域异常降温；EOF 分解的第二模态表现为典型的厄尔尼诺现象；EOF 分解的

第三模态表现为全球海域的异常升温；EOF 分解的第四模态表现为赤道附近地区的异常升温，西北太平洋海域的异常降温。EOF 分解的第五模态表现为典型的拉尼娜现象。

4. 问题二：气候变化预测

4.1 气候预测研究综述

众所周知，人们直接感受的天气现象正在发生改变，例如常见的“夏天特别热，冬天特别冷”，大量天气现象的变化说明全球气候正在悄然地发生变化。气候变化的事实说明，气候变化包括季节、年际、年代际、世纪以及更长尺度的多时间尺度的变化，引起这些不同时间尺度气候变化的原因十分复杂，不仅与气候系统外强迫因子的变化有关，还涉及气候系统内部的变化与反馈过程。简单来说，造成气候变化的原因有两种：一是自然的原因，即太阳辐射变化、火山活动等自然外强迫和气候系统内部各子系统相互作用产生的单一或耦合气候变化；二是人类活动的强迫，即人类造成的温室气体和气溶胶等排放、土地利用、植被破坏等造成的气候变化。人类活动尤其是工业化以来向大气中排放的温室气体增加，对 21 世纪全球变化产生了明显的影响。更深入地了解、研究气候的变化，有助于各国控制风险、利用不断变化的气候所带来的利益，尤其是那些最容易受气候影响的国家。

考虑到全球变暖背景下，气候变化对各地区的重要影响，气候变化中如何协调好适应与减缓的问题，如何应对好潜在风险与损失的问题，都与决策者息息相关。因而，气候预测受到越来越多的关注，同时由于不确定性也面临巨大的挑战。其不确定性主要有三个来源：一是情景的不确定性，即未来温室气体排放的不确定性；二是气候系统内部的自然变率的不确定性；三是表征气候过程的不确定性，即气候模式的结构框架不确定性。其中前两种不确定都可以人为的减小，但是气候系统内部的自然变率引起的不确定性却难以人为控制。因此，一定程度上增加了气象预测的难度。

通俗来说，现阶段的气候研究就是通过对气候变化的观测，提供气象数据，通过一定的方法进行较为准确的气象预测，为政府提供可靠的依据，使得决策者可以制定正确合理的社会经济发展规划，以从容应对气候变化对人类社会带来的冲击。现阶段，针对气象预测方面的研究，无论从实际应用还是理论研究方面都取得了快速的发展，本节将主要介绍气象预测的发展现状。

4.1.1 研究现状

气候预测根据过去气候的演变规律，推断未来某一时期内气候发展的可能趋势。人们通常会混淆天气预报和气候预测，由于人们直观感受到天气现象变化不明显，导致容易忽略长期的天气变化造成的气候变化所带来的危害。简单来说，气候预测可以看成一种长期的天气预报，而天气预报是对未来某时段内某一地区或部分空域可能呈现的天气状况所作的预测。天气预报就预报时效的长短通常分为三种：一般短期天气预报指 72 小时内的预报，中期天气预报通常指 4-10 天的天气预报，而长期天气预报短期气候预测关联性逐渐增强，通常把时效在 10 天以上的预报，一个月、一个季甚至一年的天气预报，称为长期天气预报。由于大气的可预报性一般不超过 3 周，因此，月、季、年时间尺度的天气预报称为气候预测。

值得我们关注的是，天气和气候是否存在可预报性？首先，大气系统是一个具有非线性、开放性特征的复杂系统，有多个因素共同影响，这些相互关系决定了大气系统本身具有确定性和非确定性双重特征，是一种混沌现象，其运动是非周期性的，说明大气运动缺乏可预报性，天气的变化存在不确定性；其次，由于大气的实际状态和数值模拟状态不一致，存在误差。随着时间的推移，初始误差不断变大，导致天气预报难以实现，但是气候

变化的预测不同于天气预报，后者主要依赖初值，即大气的初始状态，而前者主要是依赖于内部及外部的因素共同影响，是长期的要素影响整体气候的变化，因此气候预报可以具有可预报性；最后，一般在未来的气候预测中，实际上主要是预测由自然因素和人类活动因素对未来气候变化的可能影响，严格来说，未来气候变化情况不是真正意义上的气候预测，它只是一种气候变化的可能趋势、途径和幅度。

然后气候预测结果在什么情形下和在多大程度上是可预报的？混沌理论的研究表明，虽然在混沌系统中，个别轨迹可能是不可预报的，但整体几何形状的变化是完全可预报的。虽然气候变化本质上是一种混沌现象，表现为湍流或非周期特征，但在有些情况下也表现出相当程度的周期性或准周期性，这大大增加了气候的可预测性。目前在气候分析中广泛使用的子波分析法等。就是为揭示其不同时空尺度的周期性而进行的，地球轨道参数的米兰科维奇循环即是一个气候周期性变化的例子。

（1）国内外短期气候预测

短期气候预测研究可以追溯到 19 世纪末、20 世纪初，开展短期气候预测较早的国家有印度、前苏联、中国，接下来是美、日、英、东欧等国家。科学工作者很早就开始对气候影响因子进行探索，早在 1873 年，Blanford 就发现了印度西北 5 月大雪不利于以后季风的发展，在 1886 年他用这个因子制作了当年季风雨量的长期预报。Walker 计算了大量的相关系数，研究世界各地天气的遥相关，提出世界三大涛动的概念，并用回归方程表示三大涛动，预报印度季风的雨量。早期的短期气候预测以统计方法为主，直到 80 年代，以 ECMWF 欧洲中期数值预报中心模式为代表，各种数值模式渐渐兴起。

对气候灾害的研究已经成为一个全球性的科学问题，国际气象界制定了不少专门的国际计划对其进行研究。世界气候研究计划（WCRP）由世界气象组织与国际科学联合会联合主持，以物理气候系统为主要研究对象，是全球变化研究中开展的较早的一个计划。WCRP 组织实施的“平流层过程及其在气候中的作用”、“气候变率及其可预报性研究”、“全球能量与水循环试验”和“气候与冰冻圈”等一系列计划和行动提高了人们对气候系统的认识，推进了气候信息服务，取得了巨大成就。

同时，各国政府也在积极加强本国的气候能力建设，纷纷组织力量，开展气候预测研究工作，成立了很多国家级别的气候委员会和专门的气候研究业务机构。如美国成立了气候预测中心和具有广泛合作背景的“气候模拟实验室”，并且成立了国际气候预测研究院，加拿大、澳大利亚成立了气候中心，英国成立了哈德莱气候预测和研究中心，日本成立了气候系统研究中心以及世界上第一个厄尔尼诺预测中心等。目前气候变化的预测研究已进入全面展开并取得实质性进展的阶段，一批研究成果已开始在业务工作中得到应用。

中国是世界上开展短期气候预测业务较早的国家之一，开始于 20 世纪 30-50 年代，目前准确率在 60%-70% 之间。在我国，最早进行气候预测研究工作的是涂长望先生，他根据 Walker 的理论和方法，研究了我国气候与三大涛动的关系，并建立了中国重庆、上海、哈尔滨等站的季降水量预报的回归方程。20 世纪 50 年代开始，中央气象台开始进行月气温、月环流预报，从 1958 年正式发布汛期预报，至今开展短期气候预测服务已有 50 余年的历史。面对着国际短期气候预测研究工作的迅速发展，以及国民经济建设对短期气候预测研究日益迫切的需要，我国于年成立了国家气候中心，专门从事气候预测以及气候变化研究和分析工作、年，国家“九五”重中之重项目——《我国短期气候预测系统的研究》正式启动，为我国短期气候预测研究的发展提供了良好的机遇。

由于计算机水平的关系，曾经我国的气候预测发展很受影响。但是随着计算机水平的提高、观测数据资料的丰富、以及气象理论的精进，我国气象研究人员对于影响气候变化的物理因素及其关联关系有了更多的认识。与此同时，我国的气候预测技术在逐步的发展、提高，短期气候预测模式的物理基础和预测能力在不断的加强。2008 年，中国气象局在北

京国际新闻中心举行的新闻发布会上表示，我国的短期气候预测水平有所提高，温度和降水趋势预测已经与世界平均水平相当。

(2) 长期气候预测

长期气候预测即气候变化预估也属于气象预报的范畴。气候预估的基本原理是利用包括地球生物化学过程的耦合气候模式，在不同温室气体和气溶胶排放下，预估未来几十年或更长时期的全球或区域气候变化。但是大部分气象学的研究内容都是从气候系统内部各因素出发，从气候的成因解释气候的变化，大多都是定性的分析，难以切实为决策者提供可靠支持、难以形象客观的向大众阐明其变化规律以及变化规律带来的积极或消极影响、并且难以增加人们对气候变化的意识。

4.1.2 研究方法

(1) 全球气候系统模式研究

全球气候系统模式是进行气候变化情景预估的首要工具。针对政府间气候变化专门委员会（IPCC）评估报告的编写，世界气候研究计划（WCRP）制定的耦合模式比较计划组织了一系列模拟比较试验。众多学者利用这些模式结果，针对东亚和中国区域进行模式评估和未来气候变化预测，取得了一些很有意义的结果。

为进行变暖背景下区域尺度的气候预测，需首先对耦合模式对该区气候的模拟能力进行评估。Jiang et al. (2005) 评估了 IPCC AR3 的 7 个耦合模式在东亚区域的表现；Zhou et al. (2006) 评估了 19 个耦合气候模式对全球和中国区域表面气温的模拟能力，其结果表明，自然变率和人类活动驱动的耦合模式能较好地模拟出全球和区域尺度温度的变化特征，但对中国区域的模拟效果不如全球；模式集合模拟的中国区域 1880-1999 年温度序列与观测序列的相关系数达到 0.55，但是不能很好地再现 20 世纪后半叶以及近 10 年来显著增温趋势，亦无法再现中国中东部 27° - 26° N 间的夏季变冷趋势。Phillis and Gleckler (2006) 评估了参与 IPCC AR4 的 20 个最新全球模式对全球陆地年平均降水量的模拟能力，发现新一代全球模式对于陆地降水的模拟能力较模式的以前版本有了较大提高，模式集合模拟能力好于单个模式，但在大地形区和季风区依然存在系统偏差；Sun (2006) 研究了新一代全球模式对降水结构特征（降水频率、降水强度等）的模拟，指出最新模式依然存在降水过于频繁的缺陷；许崇海等(2007)评估了参与 IPCC AR4 的 22 个气候模式对东亚区域温度、降水的模拟能力，发现多数模式的温度模拟值偏低，降水模拟值偏高；尽管所有模式对东亚地区的气候都表现出一定的模拟能力，但各模式模拟效果差异较大。江志红等(2009)、Jiang (2012)则评估了 IPCC AR4 全球模式对中国区域极端气候指数的模拟能力，发现耦合模式对极端气候空间分布和线性趋势具有一定的模拟能力，但对年际变化模拟能力较差。田芝平等(2012)综合使用了区域均值、偏差、方差比，空间相关系数等个统计量来评估模式。张莉等(2008)以 CMAP 月平均降水资料和 ERA40 为观测基础，分析了 17 个全球海气耦合模式对东亚季风区夏季降水和环流的模拟能力。

结果表明：

① 模式基本上都能够模拟出降水由东亚东南部海洋至东亚西北部中国内陆减少的空间分布特征，部分模式能够模拟出降水的部分主要模态；

② 大部分模式基本上能够模拟出中国东部陆地降水的季节进退。但同时也存在相当的差异，这包括：多数模式普遍存在模拟降水量偏少、降水变幅偏小的缺陷；

③ 雨带的季节推进过程与观测存在一定偏差，尤其海洋上的季节进退过程模拟较差，有的模式甚至不能模拟出东亚季风区东部海洋上大致的季节进程。因此，模式对东亚季风区降水的模拟能力还是比较有限的，需要进一步改进。多模式集合的夏季环流场以偏弱为主，不利于降水的形成，这在中国东部大陆部分比较明显。另外，空气湿度模拟值偏低、

从而造成水汽输送偏弱也是导致东亚季风区夏季降水模拟偏小的原因之一。

需要指出的是，由于国家安全与社会经济可持续发展的迫切需求，年代际（10-30年）气候预测正迅速成为气候变化的一个新的优先研究领域。故耦合模式对于年代际气候变化的模拟能力也逐渐得到关注。如 Keenlyside et al. (2008) 利用 ECHAM5/MPI-OM 耦合模式，对 1966-2005 年的北大西洋邻近区域的表面温度，以及大西洋经向翻转流进行了多年代际回报检验和预测；Smith et al. (2007) 利用基于 HadCM3 模式的年代际气候预测系统，对全球表面温度的年代际变化进行了模拟。基于东亚气候在过去的近 50 年里呈现出显著的年代际变化，近年来耦合模式对年代际时间尺度气候变化的模拟评估也逐渐得到关注。如孙颖和丁一汇 (2008) 评估了 IPCC AR4 模式对东亚夏季风降水与环流年代际变化的模拟性能，发现在评估的 19 个模式中，有 9 个模式可以较好地再现中国东部地区多年平均降水场，但仅有 3 个模式（第 1 类模式）可以较好地对东亚夏季风降水的年代际变化作出模拟。

总体一般认为，气候模式区域气候温度、降水和环流皆具有一定的模拟能力，但效果不同模式间差别甚大。

（2）基于气候系统的预测方法

10-30 年时间尺度的气候预测，由于能够提供给政策制定者和有关部门与单位近期的气候变化预测，作为经济规划和社会生活安排等的参考，因此越来越受到重视。10-30 年气候预测是气候科学中一个新的领域，一般预测方法采用：一是统计预测方法，根据前期气候变化特征，利用统计关系，计算预测未来 10-30 年的气候变化；二是动力模式预测方法，利用全球或区域气候模式，在一定初始条件和外强迫作用下，预测未来 10 年的气候变化；三是动力并统计的预测方法，利用气候模式预测结果，参考统计方法建立的关系进行订正，预测未来 10 年的气候变化。气候的可预报性研究表明，从天气（日）预测到季到年际预测，主要是初值问题，而从 10 年预测到年代际或更长时期预测，主要是外强迫问题，需要指出的是，10 年预测正处在初值问题与外强迫问题交互时期。本节重点综述进行 10-30 年预测的主要气候预测方法。

（3）气候预测一般方法

气候学作为气象学的一个分支是从 19 世纪末期逐步开始形成的，气候预测技术随着数学、物理学、计算机技术及大气科学理论的发展而不断进步，其在中国的发展大致经历了经验统计分析、数理统计、物理统计、动力和统计相结合四个阶段气候预测的发展过程中，已经应用或试验的方法可以分为以下几类：

① 经验方法

在气候预测业务中，经验方法是长期使用的一种方法。通过对历史资料和天气气候的分析，利用观测数据时间序列特征作分析，进行气候预报。随着计算机技术的普及和气象资料的丰富，一些数理统计方法开始应用在气候预测方面。具体方法很多，例如周期分析、谐波分析、自回归分析、气象要素的正交函数展开等。

② 物理统计方法

物理统计方法是目前气候预测的主要统计方法。20 世纪 80 年代中期以来，气象学者开始重视物理因子的分析，利用多个物理参量作为预报因子，取得了一系列的研究成果。通常用到的预报因子有大气环流、海温场、高度场等，具体的统计方法包括：逐步回归模型、典型相关分析、经验正交函数等。

③ 动力学方法

气候动力学是大气科学与地球科学、物理学、数学以及计算机科学相互交叉的一个前沿学科。动力学方法是今后气候预报领域重点研究方法之一，动力学方法和动力统计学方法被认为是解决气候预测问题的有效方法。中尺度大气模式、MMS、ARW 等已经应用在实际业务中，优秀的数值模式对于提高预报准确率有很重要的影响。

由于影响我国气候变化的因素非常复杂，所以目前业务预报仍是以多种因素的综合分析和多种方法的综合应用为主。气象观测数据是按照一定的时间间隔做的数据记录，可以称作数据的时间序列，是很珍贵的采样数据，蕴含着丰富的气象信息，可以应用其对气候预测进行研究。

传统的时间序列分析预测模型是线性的，例如高阶自回归模型、滑动平均模型、自回归滑动平均模型，由于气候系统是一个非线性复杂的系统，线性模型的预测处理方法难以准确预测，因此引入了非线性方法进行研究。人工神经网络、混沌时间序列预测作为处理非线性问题的有效方法，将它们应用于气候预测方面，可以得到较精确的预测结果。

(4) 非线性时间序列预测方法

在气候预测方面，我们为了得到更加精确的预测结果，不能止步于传统的时间序列分析预测模型。从数据统计的角度，可以探索跨学科交叉的内容，这样也可以将深奥的气象学理论转化为易懂的数据分析理论，进而还可使数据可视化，达到气候预测的根本宗旨。

自 1927 年 Yule 发明了自回归方法，标志着现代时间序列预测的开始，在时间序列的现代预测方法中，针对单变量的预测模型主要包括：自回归 (AR, Autoregressive) 方法滑动平均 (MA, Moving Average) 方法，结合了自回归方法和滑动平均方法得到的自回归滑动平均 (ARMA, Autoregressive Moving Average) 方法等。针对多变量的预测模型主要有向量自回归 (VAR, Vector Autoregressive) 方法，向量滑动平均 (VMA, Vector Moving Average) 方法，向量自回归滑动平均 (VARMA, Vector Autoregressive Moving Average) 方法等。上述方法中大多数为线性方法，而在实际应用过程中，时间序列往往呈现非线性特性，使得上述模型的应用受到了很大的限制。针对非线性时间序列预测，研究者们提出了门限自回归 (TAR, Threshold Autoregressive) 方法，自回归条件异方差 (ARCH, Autoregressive Conditional Heteroskedasticity) 方法及其改进方法等。

上述非线性预测方法虽然在一定程度上解决了传统线性预测方法的局限性，但是它们与传统线性方法类似，均是根据专家的先验知识预先设置或构造模型结构，属于参数化模型，由于实际应用中，复杂系统由多种因素共同作用，从而产生所观测到的时间序列，根据先验知识假设的模型结构不具有自适应性，往往不能很好的反应系统的真实性质，导致模型预测性能大大降低。

4.1.3 改进措施

为了减少初始误差和模式不完善而造成的预报误差。目前的气候预测是采用多初值和多模式的集合预报方法。因而气候预测实际上是一种概率预报。

根据上节综述分析，目前利用全球气候系统模式以及各种预测方法来研究全球气候变化，已经取得了很大的进展，但由于不确定性的存在，本节对现存的问题进行以下讨论：

(1) 在实际气候预测过程中，大量的数据集在进行时间序列预测分析时，可能出现异常点或者发生数据缺失，导致预测精度低、自适应差等问题。

(2) 对比三种一般统计方法，统计预测、动力预测、统计动力预测，传统的统计气候预测则是基于古老的统计理论，并往往是事先假定了预报模式的函数形式；动力气候预测不仅计算费用高昂，而且目前的预报准确率还不能达到要求；而统计动力预测刚刚处于试验阶段。因为难免带上了许多认为主观因素。

(3) 目前，气候预测在方法上也不乏一些新技术，比如奇异谱分析、神经网络多层映射模型、单相点或多相点序列预测技术等。

4.2 基于多因素影响的非线性时间序列模型

本节内容针对非线性时间序列往往受多种“噪声”影响，数据中往往含有异常点，导

致模型预测精度低的问题。在气候预测中，我们研究气候中的温度变化特征，对于平稳变化的温度趋势，往往由于一些要素的变化出现温度的波动点，天气现象上表现为极端天气，这里包括极热以及极寒天气。我们在大量平稳的预测方法中找出，对回声状态网络的鲁棒学习方法展开研究，可以从分析不同概率分布的鲁棒性角度出发，通过分析比较，将高斯混合分布作为模型输出似然函数，并采用变分方法实现模型参数的有效学习，在贝叶斯框架下提出一种对异常点不敏感的鲁棒回声状态网络预测方法，以解决回声状态网络处理含有异常点的数据时精度差的问题。

4.2.1 气候预测模型

(1) 神经网络预测方法

众所周知，神经网络是一个具有高度非线性的大规模的动力系统。它的主要特点是能体现网络的全局作用和复杂性作大规模并行分布处理，在气候预测中，我们往往要考虑环境对系统的影响，并把现状与过去历史上曾经发生过的过程加以类比，而这两种考虑都可以与神经网络的工作方式相一致。而且神经网络的方法具有许多传统预测方法没有的优点：

- ① 容错性强，预测速度快，且可以处理信息不完全的预测问题。
- ② 由于神经网络具有一致逼近的效果，训练后的神经网络在样本上输出期望值（误差在允许范围内），在非样本点上，表现网络的联想记忆功能。
- ③ 动态自适应能力强，可适应外界新的学习样本，使网络知识不断更新，提高预测精度。
- ④ 实现非线性关系的隐式表达，不需要建立复杂的显式关系式。

其中，我们发现回声状态网络属于递归神经网络的一种，它避免了基于梯度下降方法的传统神经网络收敛速度慢，容易陷入局部最小值等缺点。训练回声状态网络本质上是一种线性回归任务，回声状态网络可以得到全局最优解，并具有更快的收敛速度。

(2) 贝叶斯回归回声状态网络预测模型

基于贝叶斯回归的回声状态网络（BESN）训练方法，不需要采用交叉验证方法便可自动实现模型参数的估计，同时能够提供模型参数的概率预测，进而得到预测结果的置信区间。但贝叶斯学习方法假设模型输出似然函数为高斯分布，这一假设使得模型对于含有异常点的时间序列非常敏感，当训练数据中含有异常点或噪声时，模型预测精度会受到很大影响，而在实际应用中，由于数据采集环境复杂，数据受多种噪声共同影响，数据中往往存在异常点。

详细地说，从上节文献综述可以发现，概率分布是一个很好表示气候变化预测不确定性的方法，他们大多运用贝叶斯多模式推断方法来预测气候变化。但是传统的贝叶斯回归方法，是假设模型误差服从零均值高斯分布，这一假设使其在含有不完整数据集，甚至异常特征值的时间序列预测中，不具有鲁棒性，稳定性差，严重影响预测精度。

我们取高斯分布的 200 个整数点的直方图分布，及其高斯分布和高斯混合分布的最大似然估计曲线，如图 4-1 所示。由于气候变化的不确定性，我们模拟突变天气的情况，将 20 个异常点加入上述数据集中产生的相应直方图分布及基于不同分布的最大似然估计曲线如图 4-2 所示，从图中可以看出，高斯分布对异常点非常敏感，从加入异常点和不加入两种情况来看，高斯混合分布不但能保存较完整的信息，避免信息损失，同时具有较强的鲁棒性，不易受异常点的影响。因此，采用高斯混合分布作为回声状态网络模型输出似然函数。

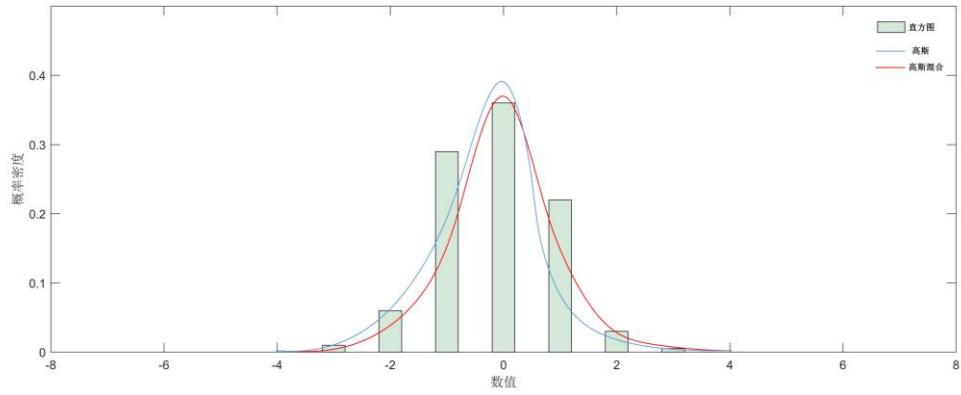


图 4-1 无异常点时各分布概率密度曲线

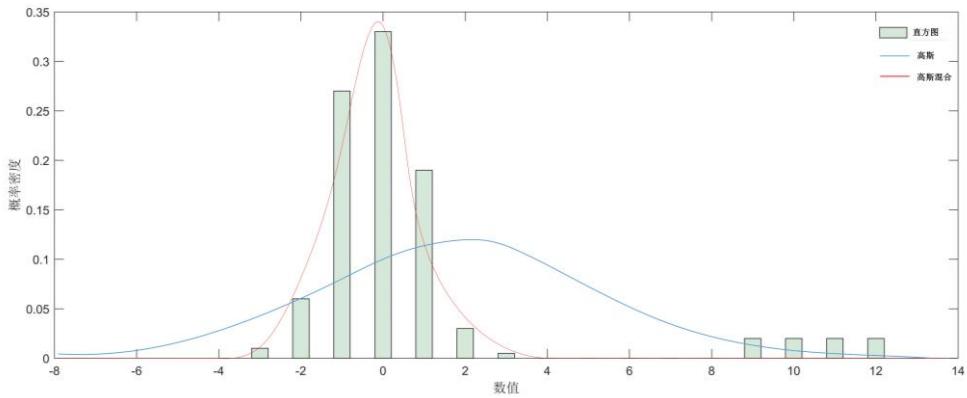


图 4-2 有异常点时各分布概率密度曲线

回声状态网络状态方程表示为如下形式：

$$\mathbf{x}(k+1) = f(\mathbf{W}_{\text{in}} \cdot \mathbf{u}(k+1) + \mathbf{W}_x \cdot \mathbf{x}(k)) \quad (4-1)$$

$$y(k+1) = \mathbf{w}^T \cdot \mathbf{x}(k+1) \quad (4-2)$$

其中， $\mathbf{u}(k+1)$ 为 $k+1$ 时刻的输入变量，若输入变量维度为 d ，则 $\mathbf{u}(k+1) \in R^d$ ，若储备池规模为 r ，则 $\mathbf{x}(k) \in R^r$ ， $\mathbf{x}(k)$ 和 $\mathbf{x}(k+1)$ 分别为 k 时刻和 $k+1$ 时刻的储备池内部状态变量， $y(k+1)$ 为 $k+1$ 时刻的输出变量，若输出为标量，则 $y(k+1) \in R$ ， $\mathbf{W}_{\text{in}} = R^{r \times d}$ 为输入权值矩阵， $\mathbf{W}_x = R^{r \times r}$ 为储备池内部连接权值矩阵，上述 \mathbf{W}_{in} 和 \mathbf{W}_x 是在网络学习训练前随机产生的。

基于贝叶斯回归的回声状态网络可以通过自动递归求得输出权值，不需要采用交叉验证方法确定正则化参数，同时能够避免伪逆法容易产生病态解的问题。该方法假设学习误差独立且服从零均值高斯分布，即输出似然函数服从如下高斯分布：

$$p(\mathbf{y} | \mathbf{w}, \beta) = (\beta / 2\pi)^{N/2} \exp \left\{ (-\beta/2) \|\mathbf{y} - \mathbf{Aw}\|^2 \right\} \quad (4-3)$$

为了便于计算， w 的先验分布如下：

$$p(\mathbf{w} | \alpha) = \frac{\alpha}{2\pi}^{r/2} \exp \left\{ -\frac{\alpha}{2} \mathbf{w}^T \mathbf{w} \right\} \quad (4-4)$$

\mathbf{w} 的相应后验分布同样为高斯分布，其均值和方差矩阵分别为 \mathbf{m}_N 和 \mathbf{S}_N ：

$$\mathbf{m}_N = \beta \mathbf{S}_N \mathbf{A}^T \mathbf{y} \quad (4-5)$$

$$\mathbf{S}_N^{-1} = \alpha \mathbf{I} + \beta \mathbf{A}^T \mathbf{A} \quad (4-6)$$

通过证据近似法确定 β 和 α 的值，计算公式如下：

$$\alpha = \frac{\gamma}{\mathbf{m}_N^T \mathbf{m}_N} \quad (4-7)$$

$$\beta = \frac{(N - \gamma)}{\|\mathbf{y} - \mathbf{A}\mathbf{m}_N\|^2} \quad (4-8)$$

$$\gamma = \sum_{i=1}^r \frac{\lambda_i}{\alpha + \lambda_i} \quad (4-9)$$

其中， λ_i 为 $\beta \mathbf{A}^T \mathbf{A}$ 的特征值，首先初始化参数 β 和 α ，再利用初始化后的 β 和 α 计算 \mathbf{m}_N 和 \mathbf{S}_N ，如公式 4-5 和公式 4-6 所示，再利用估计的 \mathbf{m}_N 和 \mathbf{S}_N ，按照公式 4-7 和 4-8 重新计算 β 和 α 的值，如此重复计算直至算法收敛。

基于贝叶斯回归的回声状态网络虽然不需要交叉验证方法就能够自动求得模型参数，但是贝叶斯回归方法假设模型误差服从零均值高斯分布，这一假设使其在含有异常点等噪声的时间序列预测中，不具有鲁棒性，稳定性差，严重影响预测精度。基于以上分析以及上述各概率分布的不同特性，选择高斯混合分布作为模型输出似然函数，并基于变分贝叶斯推理学习模型参数，得到一种既具有鲁棒性又具有自适应性的回声状态网络预测模型。

(3) 鲁棒变分回声状态网络贝叶斯推理

鲁棒回声状态网络将高斯混合分布作为模型输出似然函数，高斯混合分布是重尾分布的一种近似形式，具有对异常点不敏感的特性。对于任意一个训练样本，具体形式如下所示：

$$p(\mathbf{y}(k)) = \eta p(\mathbf{y}(k)) + (1 - \eta) p_0(\mathbf{y}(k)) \quad (4-10)$$

其中 η 是随着数据集异常点和噪声数量、比例自适应变换的参数。

$p(\mathbf{y})$ 如下：

$$p(\mathbf{y} | \mathbf{w}, \beta) = (\beta / 2\pi)^{N/2} \exp\left\{(-\beta / 2) \|\mathbf{y} - \mathbf{Aw}\|^2\right\} \quad (4-11)$$

$p_0(\mathbf{y})$ 如下：

$$p_0(\mathbf{y}) = (\beta_0 / 2\pi)^{N/2} \exp\left\{(-\beta_0 / 2) \|\mathbf{y} - \mathbf{Aw}\|^2\right\} \quad (4-12)$$

对于所有 N 个训练样本，模型输出似然函数可写成下式：

$$p(\mathbf{y} | \beta, \beta_0, \mathbf{w}, \mathbf{z}) = \prod_{k=1}^N [p_0(\mathbf{y}(k))]^{(1-z_k)} \left[\left(\frac{\beta_1}{2\pi} \right)^{1/2} \exp\left\{ -\frac{\beta_1}{2} (\mathbf{y}(k) - \mathbf{a}_k \mathbf{w})^2 \right\} \right]^{z_k} \quad (4-13)$$

其中，隐变量 \mathbf{z} 的概率分布为：

$$P(\mathbf{z} | \eta) = \prod_{k=1}^N P(z_k | \eta) \quad (4-14)$$

$$P(z_k = 0) = 1 - \eta \quad (4-15)$$

$$P(z_k = 1) = \eta \quad (4-16)$$

边缘似然函数可表示为：

$$p(\mathbf{y} | \alpha, \beta, \beta_0, \eta) = \int p(\mathbf{y} | \beta, \beta_0, \mathbf{w}, \mathbf{z}) p(\mathbf{w} | \alpha) p(\mathbf{z} | \eta) d\mathbf{w} d\mathbf{z} \quad (4-17)$$

由于上式是难以解析处理的，因此采用变分法近似推理，得到隐变量 \mathbf{z} 和网络参数 \mathbf{w} 的后验概率分布。

(4) 参数估计与模型训练

设一组训练样本输出变量为 $\mathbf{y} = [y(1), y(2), \dots, y(N)]^T$ ，模型的所有隐变量和参数表示为 \mathbf{Z} 。
 \mathbf{y} 和 \mathbf{Z} 的联合概率分布为 $p(\mathbf{y}, \mathbf{Z})$ ，变分近似推理的目标是找到后验概率分布 $p(\mathbf{Z} | \mathbf{y})$ 和模型证据 $p(\mathbf{y})$ 的一个近似解。根据变分推理，引入任意分布 $q(\mathbf{Z})$ ， $\ln p(\mathbf{y})$ 可分解为下式：

$$\ln p(\mathbf{y}) = L(q) + KL(q \| p) \quad (4-18)$$

其中

$$L(q) = \int q(\mathbf{Z}) \left\{ \ln \frac{p(\mathbf{y}, \mathbf{Z})}{q(\mathbf{Z})} \right\} d\mathbf{Z} \quad (4-19)$$

$$KL(q \| p) = - \int q(\mathbf{Z}) \ln \left\{ \frac{p(\mathbf{Z} | \mathbf{y})}{q(\mathbf{Z})} \right\} d\mathbf{Z} \quad (4-20)$$

对于离散变量，可直接将上述积分变为加和形式。上述 $q(\mathbf{Z})$ 为所有隐变量和参数 \mathbf{Z} 的联合分布， $L(q)$ 为 $\ln p(\mathbf{y})$ 的下界，根据变分推理方法，关于分布 $q(\mathbf{Z})$ 最大化下界 $L(q)$ 即是最小化 $KL(q \| p)$ ，若 $q(\mathbf{Z})$ 可取任意分布，当 $KL(q \| p)$ 为 0 时， $L(q)$ 达到最大值，但是 $q(\mathbf{Z})$ 为由一组参数和隐变量决定的特定分布，不具有任意性。根据文献，当 $q(\mathbf{Z})$ 可分离时， $KL(q \| p)$ 达到最小值， $q(\mathbf{Z})$ 可分离即是满足下式：

$$q(\mathbf{Z}) = \prod_{i=1}^s q_i(Z_i) \quad (4-21)$$

$$\begin{aligned} L(q) &= \int \prod_i q_i \left\{ \ln p(\mathbf{y}, \mathbf{Z}) - \sum_i \ln q_i \right\} d\mathbf{Z} \\ &= \int \prod_i q_i \left\{ \ln p(\mathbf{y}, \mathbf{Z}) - \prod_{i \neq j} q_i d\mathbf{Z}_i \right\} d\mathbf{Z}_j - \int q_j \ln q_j d\mathbf{Z}_j + const \\ &= \int q_j \ln \tilde{p}(\mathbf{y}, \mathbf{Z}_j) d\mathbf{Z}_j - \int q_j \ln q_j d\mathbf{Z}_j + const \end{aligned} \quad (4-22)$$

其中，

$$\ln \tilde{p}(\mathbf{y}, \mathbf{Z}_j) = E_{i \neq j} [\ln p(\mathbf{y}, \mathbf{Z})] \quad (4-23)$$

const 表示一个常数, $E_{i \neq j} [\ln p(\mathbf{y}, \mathbf{Z})]$ 表示 $\ln p(\mathbf{y}, \mathbf{Z})$ 关于所有 $i \neq j$ 的变量 \mathbf{Z}_i 为 q 分布的期望, 即

$$E_{i \neq j} [\ln p(\mathbf{y}, \mathbf{Z})] = \int \ln p(\mathbf{y}, \mathbf{Z}) \prod_{i \neq j} q_i d\mathbf{Z}_i \quad (4-24)$$

式 4-22 为 $q_j(\mathbf{Z}_j)$ 和 $\tilde{p}(\mathbf{y}, \mathbf{Z}_j)$ 之间的负的 KL 散度, 最大化公式 4-22 等价于最小化该 KL 散度, $\tilde{p}(\mathbf{y}, \mathbf{Z}_j) = q_j(\mathbf{Z}_j)$ 时, $L(q)$ 达到最大值, 整理得,

$$\ln q_j(\mathbf{Z}_j) = E_{i \neq j} [\ln p(\mathbf{y}, \mathbf{Z})] + const \quad (4-25)$$

根据式 4-21 和 4-25 和贝叶斯理论可得,

$$\ln q_{\mathbf{w}}(\mathbf{w}) = E_{\mathbf{z}} [\ln p(\mathbf{y}, \mathbf{w}, \mathbf{z})] + const \quad (4-26)$$

$$\ln q_{\mathbf{z}}(\mathbf{z}) = E_{\mathbf{w}} [\ln p(\mathbf{y}, \mathbf{w}, \mathbf{z})] + const \quad (4-27)$$

$$\ln p(\mathbf{y}, \mathbf{w}, \mathbf{z}) = \ln p(\mathbf{y} | \mathbf{w}, \mathbf{z}) + \ln p(\mathbf{w} | \alpha) + \ln(\mathbf{z} | \eta) \quad (4-28)$$

最终计算得到 $\ln p(\mathbf{y}, \mathbf{w}, \mathbf{z})$, 再基于公式 4-25 和 4-26, 求得模型参数 \mathbf{w} 的近似后验概率分布为高斯分布, 其协方差矩阵和均值分别为 μ 和 Σ :

$$\Sigma = \left(\sum_{k=1}^N [\beta_0(1 - E_z(z_k)) + \beta E_z(z_k)] \mathbf{a}_k^T \mathbf{a}_k + diag(\alpha_h) \right)^{-1} \quad (4-29)$$

其中,

$$diag(\alpha_h) = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_r \end{pmatrix} \quad (4-30)$$

$$\mu = \Sigma \left(\sum_{k=1}^N [\beta_0(1 - E_z(z_k)) + \beta E_z(z_k)] y(k) \mathbf{a}_k^T \right) \quad (4-31)$$

$E_{q_{z_k}}(z_k)$ 是 z_k 关于分布 $q_{z_k}(z_k)$ 的期望, 隐变量 \mathbf{z} 的概率分布为:

$$q_{\mathbf{z}}(\mathbf{z}) = \prod_{k=1}^N q_{z_k}(z_k) \quad (4-32)$$

$$q_{z_k} = (b_k^{z_k} c_k^{1-z_k}) / (b_k + c_k) \quad (4-33)$$

其中,

$$b_k = \beta^{1/2} \eta \exp(-\beta d_k / 2) \quad (4-34)$$

$$c_k = \beta_0^{1/2} (1 - \eta) \exp(-\beta_0 d_k / 2) \quad (4-35)$$

$$d_k = (y(k) - \mathbf{a}_k \mu)^2 + \mathbf{a}_k \Sigma \mathbf{a}_k^T \quad (4-36)$$

根据变分贝叶斯方法, 参数更新公式如下:

$$\alpha_h^{new} = \frac{\delta_h}{\mu_h^2} \quad (4-37)$$

$$\delta_h = 1 - \alpha_h \Sigma_{hh} \quad (4-38)$$

$$\beta^{new} = \frac{\sum_{k=1}^N E_z(z_k)}{\sum_{k=1}^N E_z(z_k) d_k} \quad (4-39)$$

$$\eta^{new} = \frac{1}{N} \sum_{k=1}^N E_z(z_k) \quad (4-40)$$

其中， μ_h 是 μ 的第 h 个元素， Σ_{hh} 是矩阵 Σ 对角元素上的第 h 个元素。

4.2.2 多因素作用下的鲁棒变分回声状态网络预测方法

由于单一鲁棒回声状态网络对输入时间序列的动态特性挖掘能力有限，为进一步提高单一模型预测进给系统时间序列的预测精度，不同于传统的预测方法利用已有的单一温度数据，根据时间序列描述其过去的变化规律以得出未来的预测趋势。

为了进一步挖掘已有的数据，根据资料收集，我们找到影响气候特征温度的因素，包括地球的吸热和散热、海洋的温度变化等，目的是研究气候变化过程中这些因素对每个温度点的影响状态，再对未来的温度状态进行预测。

由于影响气候变化的因素有很多，并且输入变量直接也存在相关关系，使其提供的信息部分发生交叉和重叠，影响数据的正确分析，因此我们利用主成分分析法进行降维，对气候预测的输入变量进行预处理。

利用主成分分析法对回声状态网络的输入变量进行预处理，从而建立多因素时间序列预测模型的主要步骤为：

- (1) 消除每个因素变量的量纲不同对主成分分析和回声状态网络建模所带来的影响，对数据进行标准化处理，得到归一化样本数据；
- (2) 求出影响因素标准化数据的样本相关矩阵，进一步求出样本相关矩阵的特征值、对应的特征向量和贡献率，特征值按降序排列 $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots \geq \lambda_n \geq 0$ (n 为影响因素变量个数)；
- (3) 选取主成分个数 p ，并计算主成分值；
- (4) 进行回声状态网络学习。

4.2.3 算法流程

根据多因素作用下的鲁棒变分回声状态网络预测模型，首先利用主成分分析将输入变量进行重构，得到一组重构后的多元耦合时间序列，将新的时间序列输入到回声状态网络，通过储备池进行高维空间映射，舍弃暂态过程，基于映射后的高维空间数据变量，进行鲁棒贝叶斯推理，再利用变分近似推理方法求得网络输出权值 \mathbf{w} 。

模型具体实现过程如下：

第一步：若出现多因素影响的情况下，采用主成分分析法对输入变量进行降维，达到时间序列重构效果；

第二步：随机初始化储备池输入权值矩阵 \mathbf{w}_{in} 和内部连接权值矩阵 \mathbf{w}_x ，并选择适当的储备池规模 r 及内部连接权值矩阵的稀疏度和谱半径，将第一步得到的新时间序列作为储备池输入变量，映射到高维空间，舍弃暂态过程，得到高维序列矩阵 \mathbf{A} 。由回声状态网

络贝叶斯推理及鲁棒回声状态网络贝叶斯推理过程，将储备池输出矩阵 \mathbf{w} 作为待估参数，设其先验概率分布为高斯混合分布，储备池目标输出似然函数，隐变量 \mathbf{z} 的概率分布为式。

第三步：基于变分推理方法，对鲁棒回声状态网络输出权值进行估计。

(1) 初始化 $\alpha_h (h=1, 2, \dots, r)$ 、 β 、 β_0 、 η 及 $q_{z_k}(z_k)$

(2) 采用上述初始化后的参数值，利用式计算协方差矩阵 Σ ，再利用求得的 Σ ，通过式计算均值 μ 。获得 μ 后再继续利用式计算协方差矩阵 Σ ，如此循环 v_0 次， v_0 称为子更新次数。

(3) 利用式实现参数 α_h 、 β 及 η 的更新，参数更新后再返回到上述(2)步，更新 Σ 和 μ ，如此循环 v_1 次， v_1 称为主更新次数，直至算法收敛。

第四步：将第三步估计得到的 μ 作为鲁棒回声状态网络的输出权值 \mathbf{w} ，当有新样本需要预测时，将新样本通过第一步和第二步进行主成分分析和高维空间映射，并利用回声状态网络方程得到其预测值。

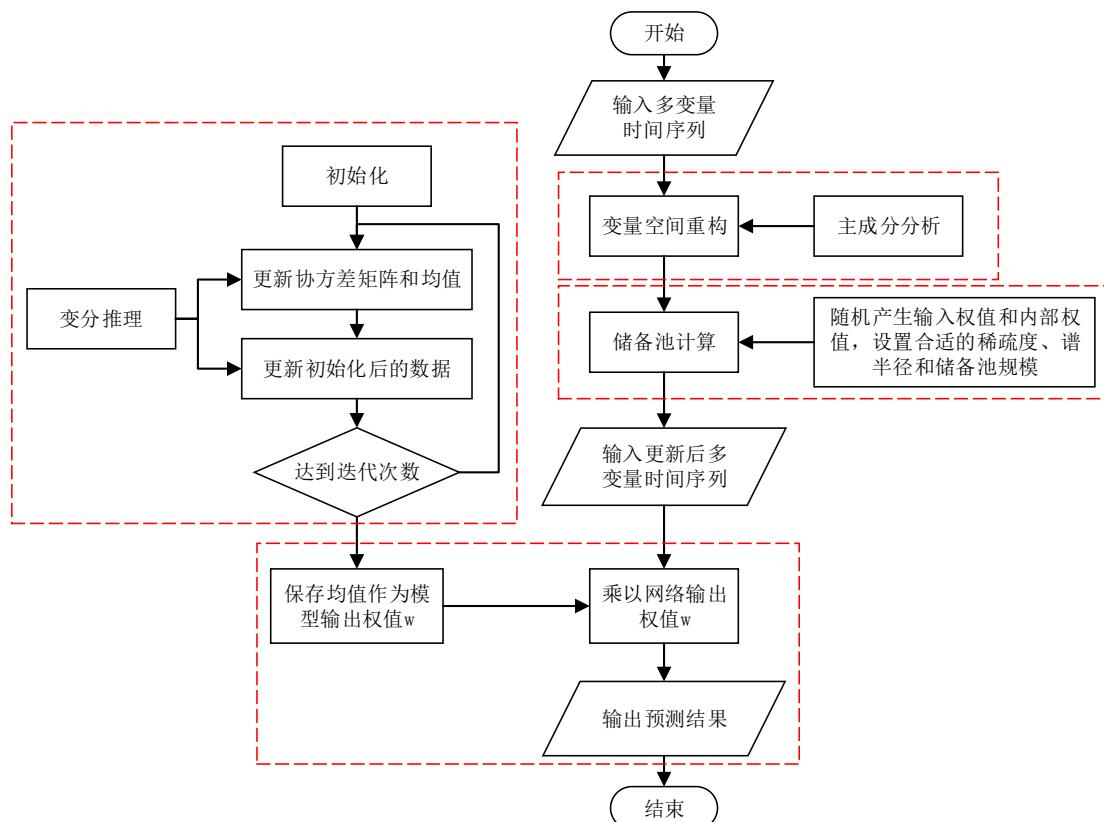


图 4-3 多因素影响下的鲁棒回声状态网络预测方法流程图

4.3 计算结果

依据题目意思，需要对未来 25 年的气候变化进行预测，我们基于已有的温度时间序列做预测，不同于传统时间预测，根据已有的历史数据对未来进行简单预测。针对气候预测的需求以及一般方法的局限性，我们利用上述所提的多因素影响下的鲁棒回声状态网络预测方法对全球温度进行预测，以得到一个全球气候变化的趋势，其中我们主要以地球吸热、散热、海洋温度变化为要素分析，关键步骤及优势如表 4-1 所示。

表 4-1 模型优势分析

回声状态网络	针对数据缺失	容错性强、预测精度高、自适应强
变分推理	针对异常点的出现	鲁棒性
主成分分析	针对多要素影响	提高精度、消除要素重叠影响

4.3.1 验证模型可行性

以 1951-2000 年的相关数据带入模型模拟出全球平均气温变化值，并与实际观测值进行比较。

根据上节得出的回声状态网络预测模型我们可以根据地球的吸热散热和海洋温度变化（太阳辐射强度）、温室气体浓度要素对温度进行预测。

我们模拟出 1951-2000 年全球气温变化如图所示。

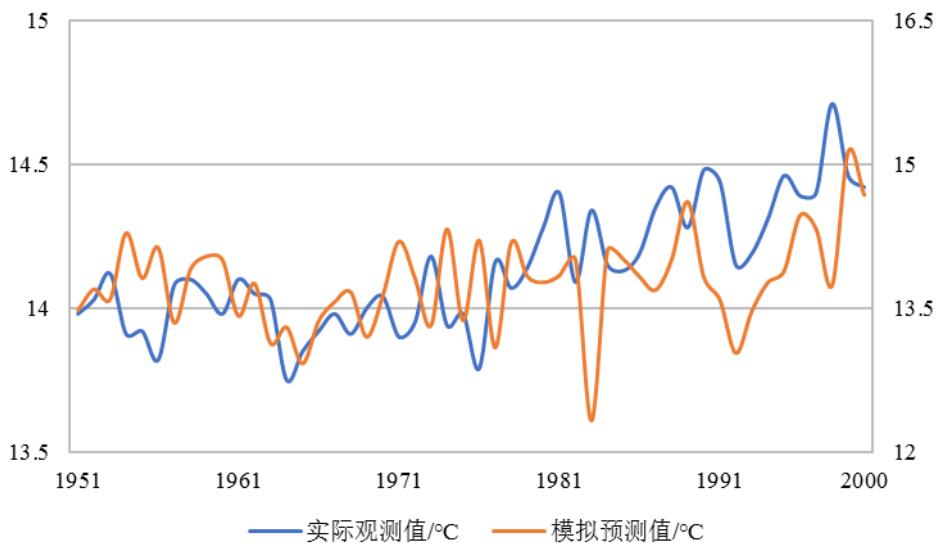


图 4-4 1951-2000 年全球平均气温模拟值与实际观测值关系图

由图可以看出，模拟曲线的年际变化要比实际观测值大一些，另外我们还可以发现一个很有趣的现象，就是全球平均气温模拟值的大部分波峰或波谷恰好是对应实际观测值的波谷或波峰。而 20 世纪 90 年代初两者却同时达到一个波谷。

模拟值是根据每年太阳辐射和海洋温度变化确定的，根据目前研究确认模型中太阳辐射与模拟值之间是一个简单的线性关系，那么也就是说计算得的太阳辐射影响的温度序列的波峰或波谷应该恰好对应实际观测的平均气温的波峰或波谷。

但是根据模型建立的原理，我们除了加入太阳辐射，还考虑海洋温度变化对全球气温的影响。分析后认为第一个现象很可能是由于地球上存在大面积的海洋。海水的比热容较大，对气温变化有很好的缓冲作用。

因此我们验证的模型能有效反映实际情况。

4.3.2 全球气温变化预测

(1) 主要要素的预测

① 太阳辐射预测

根据太阳变化规律，以及 1881-2000 年太阳辐射强度变化，可知自上个世纪年代以来太阳辐射强迫基本呈稳定的以 11 年为周期的周期性变化，在本次研究中我们假设这种变化

将继续延续下去。以 1957-2000 年 44 年中（4 个太阳活动周期）太阳辐射的平均周期变化规律作为未来近 25 年太阳辐射的变化规律。

② 温室气体

根据资料显示上个世纪 60 年代以前主要温室气体产生的总辐射强迫增长速度缓慢，一直处于一个较低的水平。60 年代至 90 年代初期，总辐射强迫快速增长，之后逐渐趋于稳定，如图所示。

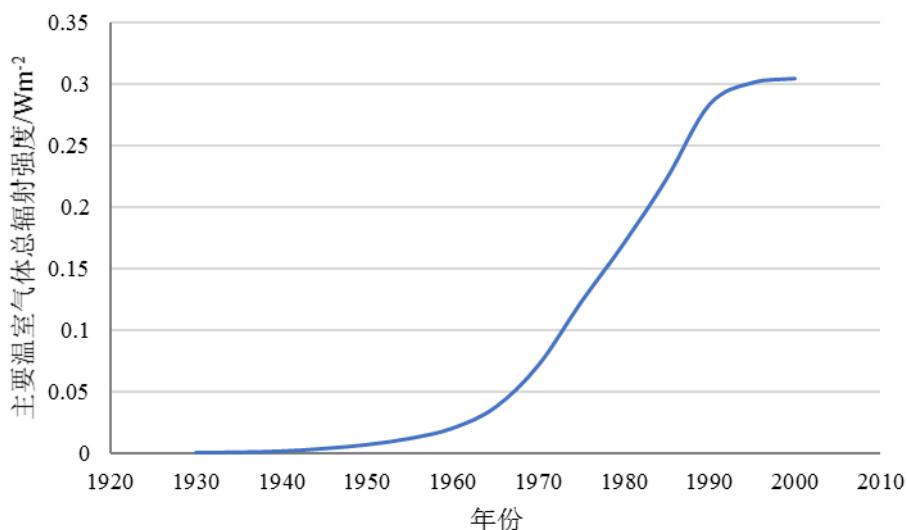


图 4-5 1930-2000 年全球主要温室气体的总辐射强度变化图

由于温室气体在大气中的生命周期都较长，因此在本次研究中假设未来 25 年，主要温室气体产生的辐射强迫保持稳定不变。

(2) 全球气温变化预测结果

根据各要素的预测值以及上节模型，算得 2020-2045 年的全球平均气温，具体数据见附表 2，得到预测数值曲线如图所示。

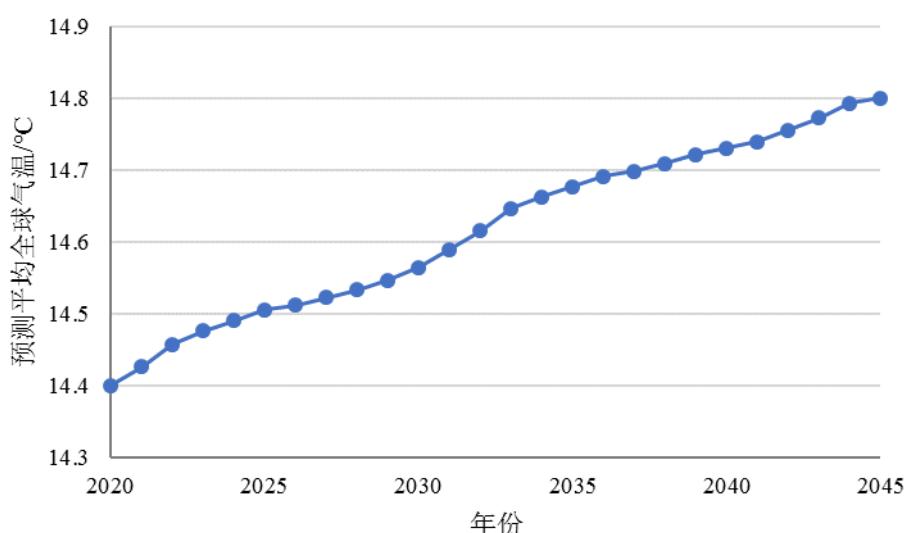


图 4-6 2020-2045 年全球平均气温预测变化趋势图

从图 4-6 中可以看出，2020-2045 年全球气温将升高 0.4°C；期间全球平均气温预测值

总体上将不断升高，但升高的趋势逐渐放缓。根据资料显示，这是由于地球的长波辐射与自身的温度的4次方成正比，也就是说，随着全球平均气温的增加，地球的长波辐射的增长速度更加快，并以此作为负反馈来抑制全球平均气温的升高，因此全球平均气温的增长速度将逐渐放缓。

从图中还可以看出，全球平均气温变化的增长速度与太阳活动周期一致，呈以11年为周期的变化。这是由于除了太阳辐射强迫是呈周期性变化外，其余因子在预测中都大致假定成线性增长或者不变。

为了体现模型具有良好的精度、鲁棒性以及自适应性，我们在训练样本时，采用不同的模型进给速度得到四组不同的训练样本。在将四组数据分别采用传统的预测方法进行预测，其中包括多元线性回归、贝叶斯回归和支持向量机。

采用均方差误差定量评价模型的性能，得到不同的精度结果。

表 4-2 不同模型进给系统时间序列预测结果

模型	F=2000	F=3000	F=4000	F=5000
多元线性回归	2.5100	2.5703	3.3394	4.9403
贝叶斯回归	1.9718	2.1086	2.8835	4.2720
支持向量机	0.3266	0.3692	0.6055	0.7087
RVESN	0.2800	0.3068	0.5365	0.6563

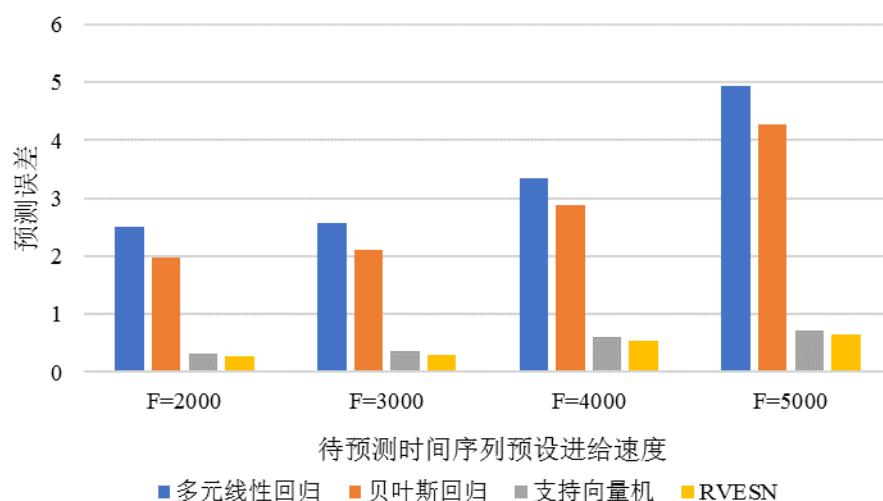


图 4-7 不同预测方法精度对比图

从图中能够更清楚的看出，所提方法和支持向量机方法都具有一定的鲁棒性，而本研究方法相对于支持向量机具有更高的预测精度。从图中可以看出，由于气候变化的不确定性导致出现异常点，多元线性回归模型和贝叶斯回归模型两种经典模型预测精度都受到了很大的影响，而本研究的模型具有更好的自适应性，对于进给系统时间序列具有较好的预测性能。

5. 问题三：“极寒天气”与“全球变暖”

5.1 问题描述

气候变化是指特定地区的气候平均状态随时间的变化（一般为年代际尺度），这与天气的短时变化（几分钟到几天）是不一样的。气候变化可能由多种因素造成，包括太阳黑子等太阳活动引起的地表吸收的净辐射变化、地球轨道的变化、板块运动或大规模的火山活动，以及大气和海洋环流变化等自然因素。人类活动，特别是工业革命以来的人类活动，包括化石燃料燃烧和其他工业活动所造成的二氧化碳等温室气体的排放、土地利用变化、大规模城市化等，也会引起气候的变化。温室气体能够吸收地球表面的辐射能量，阻止红外辐射直接进入太空，由此造成的“自然温室效应”会使地球温度升高。许多研究表明气候变化也称全球变暖。以全球变暖为主要特征的气候变化不仅对人类生存环境造成严重的影响，并可能进一步导致水资源、生态系统状况发生变化，造成旱涝等自然灾害的频发，对工农业生产，社会经济发展、人类健康和政治格局等产生深远的影响。气候变化及其影响已成为当前全球性的重大科学问题，并受到各国政府和公众的关注。

极端气候事件，主要包括极端温度、极端降水两个方面。极端气候事件是全球受关注度最高、影响最为巨大的自然灾害之一，是学术界研究的一个热点问题。极端气候事件的发生越来越频繁，而生态系统对气候变化的适应能力在不断减弱，这使得极端气候事件的危害性越来越明显。极端天气、气候事件的频繁发生给人类的生产生活、自然环境、生态系统、经济发展及生命财产安全都造成了严重的威胁与破坏。

极寒天气属于极端气候事件，根据气象专业制定的“寒冷程度等级表”，气温从-40℃以下至9.9℃，由低到高共分为八级。其中一级为“极寒”，代表-40℃以下的天气。

“极寒天气”是某地的天气现象，这种极端气象的出现，与全球的气候变化有无关系？全球变暖和局地极寒现象的出现之间是否矛盾？

5.2 分析思路

针对问题三的子问题一，我们介绍了两种方法探究“极寒天气”与气候变化有无关系：首先是从地理学的角度，定性分析极端气象的出现的影响因素及相关过程，分析两者有无关系；然后借助极值理论，以加拿大NUNAVUT地区为例，分析该地区近20年来极寒天气出现的变化特征，讨论某地极寒天气的出现与该地区气候变化的关系。

针对问题三的子问题二我们构建了分层的贝叶斯时空模型，进一步分析某地的极寒天气的出现与全球气候变化之间的关系，证明全球变暖和局地极寒现象的出现之间是否矛盾的问题。

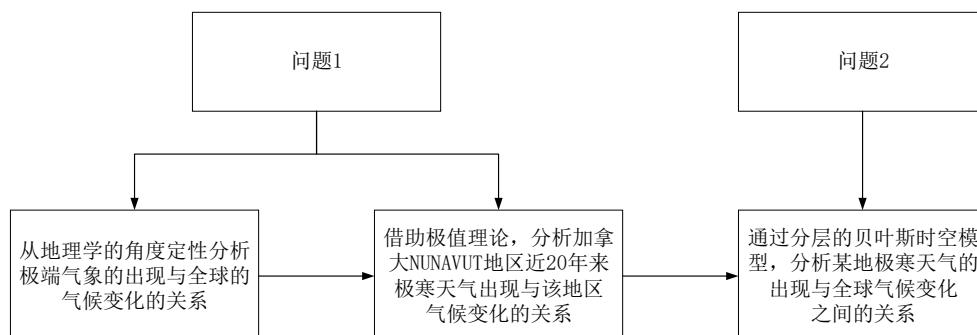


图 5-1 具体流程图

通过以上这些方面，全面的描述和分析极端气象的出现与全球的气候变化之间的关系。

5.3 有无关系探究：方法 1—地理特征

近年来，全球各地频频出现极寒天气事件，而且有愈演愈烈的趋势。根据气象专业制定的“寒冷程度等级表”，气温从-40℃以下至 9.9℃，由低到高共分为八级。其中一级为“极寒”，代表-40℃以下的天气。这些极寒天气的出现，与全球气候变化也就是全球变暖息息相关。

描绘地球的最好方式是把它看成绕其轴旋转的球体，但是有两种附加效应：大气和海洋。随着地球绕地轴旋转，我们感受到昼夜交替，白天在太阳光照射下感觉天气变暖，夜间在黑暗中感觉天气变冷，因为地球把其储存的热量辐射到太空深处。当我们所处的半球朝着太阳倾斜时，我们就经历了夏季；当我们所处的半球背离太阳倾斜时，我们就经历了冬季。海洋储存了巨大热量，由洋流把热量从一个地方带到另一个地方。但是，在我们目前经历的这些极端天气事件中，大气才是最重要的影响因素。

在任何自转的行星上，都会产生叫作“盛行风”（Prevailing winds）的效应。随着自转，地球一般会经历 3 种不同的风，通常分别局限在 3 个不同纬度地带：

纬度 0° 到 30° 地带：盛行信风，自东向西吹，汇聚于赤道。

纬度 30° 到 60° 地带：盛行西风，自西向东吹，上升至北极圈或者下沉至南极圈。

纬度 60° 到 90° 地带：极地环流，通常局限在地球纬度最高的地区。

虽然纬度带不同，这种现象对于大多数快速自转、具有大气的行星都很常见，包括金星、火星、木星和土星。不过，地球有一点特别。由于地球大气层很稀薄、地球的大幅轴向倾斜、极地上空的云层遮蔽和反射行为以及一系列其他因素，地球赤道和两极之间的温差极大。因此，在两极存在持续的大范围低压带，以气旋的方式自西向东旋转（在北极是逆时针方向旋转，在南极是顺时针方向旋转）。这两个低压带称为“极涡”，它们均始于大气层上空数英里处，然后充分扩展到平流层。

在极涡下面，存在大量密度大的冷空气环绕着两个极点。正常情况下，这些极涡足够稳定，因为温差和气压差大到足以让极涡一年到头都留在原地。当极涡达到最强时，就变成单个环流，空气固定在一定区域。当极涡变弱时，它们会分裂成两个或多个环流，并且开始偏离极点。当极涡变得极弱时，它们会分裂成碎片，其中一些低压、低温的空气会开始与极区之外气压和温度更高的空气发生相互作用。

北极地区大部分是海洋，被欧亚、北美大陆所包围，就地、气系统而言，极区接受的太阳短波辐射少，地表长波辐射强，导致极区有大量的热量丧失，因而极区空气极端寒冷，高空大气从中纬度流向极地之后堆积下沉形成下层的极地高压，气压梯度力作用下，极区近地面强冷空气南移，给中纬度地区天气气候带来极大影响。也就是当北极的极涡变得极弱时，地球上纬度地区的高压带（西风带所在之处）会向极点推进，替换冷空气。这导致极涡更加向南移动。而且，急流扭曲并向人口稠密的南纬地区偏离。当来自极区的干冷空气遇到中纬度地区的湿热空气，就会出现剧烈的天气变化，也就是极寒天气。

导致极涡分裂的现象称为“平流层突然变暖”，平流层上层的空气在短短几天之内变暖大约 30~50℃。地表温度越高，尤其是在北美和欧亚大陆北部地区，就会允许越多的热量传输到北极平流层。变暖的地球使得平流层突然变暖事件更大可能和更频繁。而这些事件会使极涡变得不稳定，把冷空气带到中纬度地区，导致我们目前正在经历的极端天气。

随着地球持续变暖，这些主要地区的积雪和海冰将会减少，从而改变极涡边界地区的气压和气温梯度。在极端情况下，造成的结果是极涡变弱乃至崩溃。

5.4 有无关系探究：方法 2—极值理论

5.4.1 相关理论

(1) 泊松点过程 (Poisson Point Process) 理论

设在 t 时刻某事件(如极寒天气)发生的次数为 $N(t)$ ，当满足下面的两个条件时， $N(t)$ 称为泊松过程：

① 在不同的时间区间内，事件发生的次数是相互独立的（这一假设在实际中一般都是成立的）；

② 在时间区间 $[s, t]$ 内，事件发生的次数服从泊松分布，即

$$\Pr(N(t) - N(s) = k) = e^{-\mu} \frac{\mu^k}{k!}, k = 0, 1, \dots \quad (5-1)$$

其中期望为 $\mu = \int_s^t \lambda(x) dx$

$\lambda(x)$ 是泊松过程的强度函数 (intensity function)。强度函数 $\lambda(x)$ 能够反映出在点 x 该事件发生可能性， $\lambda(x)$ 越大事件发生可能性就越大。若 $N(t)$ 是泊松过程，那么时间发生的具体时刻就构成了泊松点过程。当 $\lambda(x)$ 不随 x 的变化而变化，即 $\lambda(x)$ 是常数时，泊松点过程为齐次泊松点过程，当 $\lambda(x)$ 不为常数时，泊松点过程为非齐次泊松点过程。

(2) 泊松回归定理

在对极端气候事件进行趋势分析时，之前的绝大多数研究都是采取普通最小二乘法、线性拟合等分析方法。然而在分析极端气候事件及连续极端气候事件发生天数的变化趋势特征时，我们应当注意，这里反应变量（极端气候事件及连续极端气候事件的发生天数）是计数变量，并且服从泊松分布：

$$P(y) = \frac{\lambda^y}{y!} e^{-\lambda} \quad (5-2)$$

破坏了最小二乘法的基本假定条件，因而不适宜应用最小二乘法来进行分析。本题采用的是泊松回归的分析方法。泊松回归是一种广义线性回归，它可以为计数因变量模型提供精确的回归结果。此处所建立的泊松回归模型如下：

$$\ln(Y) = \beta_0 + \beta_1 X^2 + \beta_2 X^2 \quad (5-3)$$

模型也可以写作：

$$\ln(\lambda) = \beta_0 + \beta_1 X^2 + \beta_2 X^2 \quad (5-4)$$

其中 λ 是 Y 的数学期望， β_1, β_2 分别是一次项和二次项的回归系数。

(3) 广义极值分布 (GEV) 定理

广义极值 (GEV) 分布主要应用于区组模型。令 X_1, X_2, \dots, X_n 是一列独立同分布的随机变量，在实际应用中 X_1, X_2, \dots, X_n ，表示一个随机过程，而 $M_n = \max\{X_1, X_2, \dots, X_n\}$ 则表示这个随机过程的最大值。如果存在着常数序列 $\{a_n > 0\}$ 和 $\{b_n\}$ 使得：

$$\Pr\{(M_n - b_n)/a_n \leq z\} \rightarrow G(z) \quad (\text{当 } n \rightarrow \infty \text{ 时}) \quad (5-5)$$

这里 $G(z)$ 是非退化分布，则此极限分布 $G(z)$ 必然是下列分布族中的一种：

① Gumbel 类型:

$$G(z) = \exp\left\{-\exp\left[-\left(\frac{z-b}{a}\right)\right]\right\} \quad (5-6)$$

② Frechet 类型:

$$G(z) = \begin{cases} 0, z \leq b \\ \exp\left\{-\left(\frac{z-b}{a}\right)^{-\alpha}\right\}, z > b \end{cases} \quad (5-7)$$

③ Weibull 类型:

$$G(z) = \begin{cases} \exp\left\{-\left[\left(\frac{z-b}{a}\right)^\alpha\right]\right\}, z < b \\ 1, z \geq b \end{cases} \quad (5-8)$$

这三种类型的分布可以归结为统一的形式，即广义极值分布 (GEV) 族:

$$G(z) = \exp\left\{-[1 + \xi(\frac{z-\mu}{\sigma})]^{-1/\xi}\right\} \quad (5-9)$$

其中定义域为: $\{z : 1 + \xi(z - \mu) / \sigma > 0\}, -\infty < \mu < \infty, \sigma > 0, -\infty < \xi < \infty$ 。

(4) 广义帕累托分布 (GPD)

GEV 分布基于区组模型。在每个区组里找到最大值，进而探讨这些区组最大值的渐进分布。为了尽可能的多使用有用的数据能够减少我们的估计偏差，从而我们需要考虑基于过顶点模型 (POT) 的广义帕累托分布 (GPD)。

对于独立同分布的随机变量序列

$$X_1, X_2, \dots, X_n \quad (5-10)$$

其区组最大值

$$M_n = \max\{X_1, X_2, \dots, X_n\} \quad (5-11)$$

若对于充分大的 n ，

$$\Pr\{M_n \leq z\} \rightarrow G(z) \quad (5-12)$$

其中 $G(z)$ 满足 (5-9) 式，则对于足够大的 u ， $X - u$ 的条件分布近似为:

$$\Pr\{X - u \leq y | X > u\} = H(y) = 1 - (1 + \frac{\xi y}{\sigma})^{-1/\xi} \quad (5-13)$$

定义域为: $\{y : y > 0 \text{ 且 } (1 + \xi y / \tilde{\sigma}) > 0\}$ ， 其中 $\tilde{\sigma} = \sigma + \xi(u - \mu)$ 。

(5) Mann-Kendall 检验

Mann-Kendall 的检验方法是非参数方法。非参数检验方法亦称为无分布检验，其优点是不需要样本遵从一定的分布，也不受少数异常值的干扰，更适用于类型变量和顺序变量，计算也比较简便。具体方法如下。

对于具有 n 个样本量的时间序列 X ，构造秩序列:

$$s_k = \sum_{i=1}^k r_i \quad r_i = \begin{cases} 1 & x_i > x_j \\ 0 & \text{else} \end{cases} \quad j = 1, 2, \dots, i \quad (5-14)$$

可见，秩序列 s_k 是第 i 时刻数值大于 j 时刻数值个数的累计数。在时间序列随即独立的假定下，定义统计量

$$UF_k = \frac{s_k - E(s_k)}{\sqrt{Var(s_k)}} \quad k = 1, 2, \dots, n \quad (5-15)$$

式中 $UF_1=0, E(s_k), Var(s_k)$ 是累计数 s_k 的均值和方差，在 x_1, x_2, \dots, x_n 相互独立且有相同连续分布时，它们可由下式给出：

$$E(s_k) = \frac{n(n+1)}{4} \quad Var(s_k) = \frac{n(n-1)(2n+5)}{72} \quad (5-16)$$

UF_i 为标准正态分布，它是按时间序列的顺序 x_1, x_2, \dots, x_n 计算出来的统计量序列，给定显著性水平 α ，查正态分布表，若 $|UF_i| > U_a$ ，则表明序列存在明显的趋势变化。

按时间序列的逆序 x_n, x_{n-1}, \dots, x_1

再重复上述过程，同时使 $UB_k = -UF_k, k = n, n-1, \dots, 1 \quad UB_1=0$ 。

5.4.2 数据来源和处理

加拿大总面积 998 万平方公里，位居世界第二位，其中陆地面积 909 万平方公里，淡水覆盖面积 89 万平方公里。全国分 10 省 3 地区。10 省为不列颠哥伦比亚、阿尔伯塔、萨斯喀彻温、曼尼托巴、安大略、魁北克、新不伦瑞克、诺瓦斯科舍、爱德华王子岛、纽芬兰和拉布拉多，3 地区为育空、西北、努纳武特。其中努纳武特地区，位于加拿大东部北极地带。努纳武特地区至少有一半时间是白茫茫的冰天雪地，温度有时候降至零下 40 度，也就是会出现极寒天气，因此选择该地区作为研究对象，具有合理性和一定的代表性。

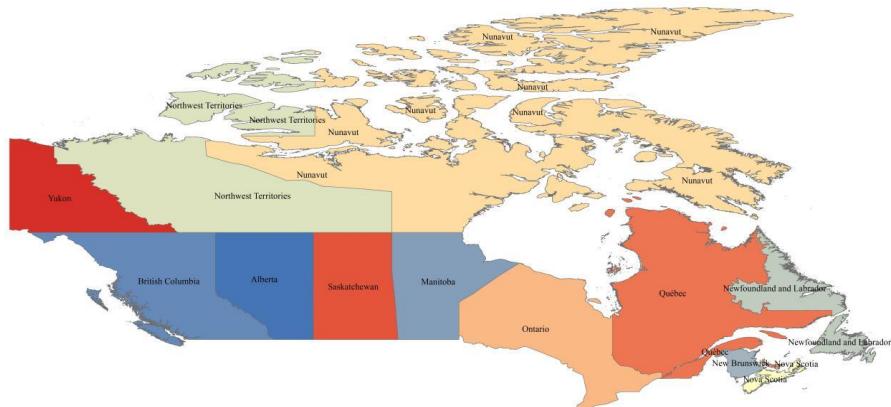


图 5-2 加拿大行政区划图

此处以第一题中加拿大全国 10 省 3 地区 2000 年以后近 20 年逐日气温资料为基础，首先统计出监测到过极寒天气的区域，结果如下图所示：

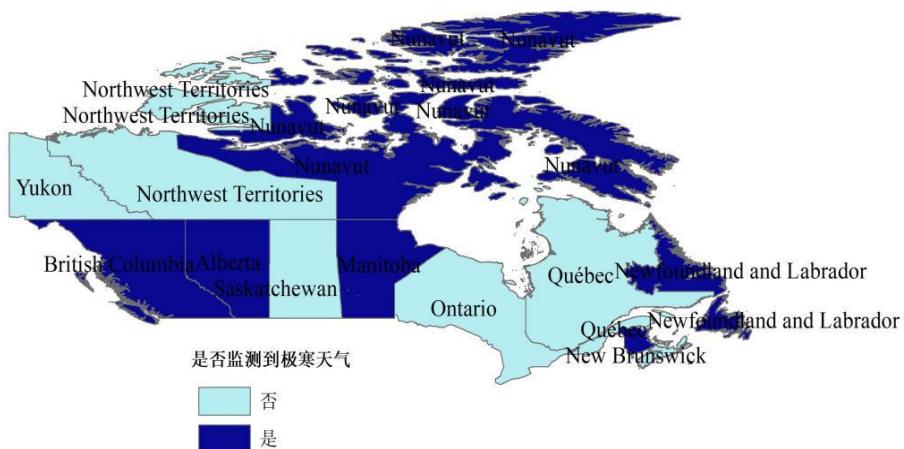


图 5-3 极寒天气区域分布图

然后对 NUNAVUT (简称 NU) 地区近 20 年逐日气温资料进行筛选，得到每年最低温度低于零下 40 摄氏度（即出现极寒天气）的相关统计数据，并做出该地区极寒天气年出现天数图：

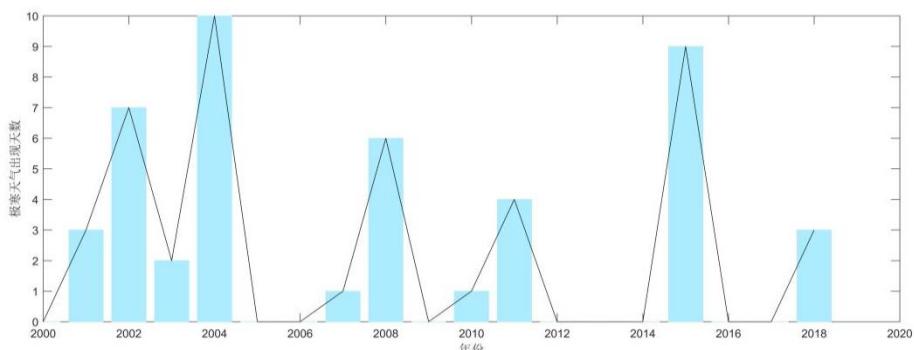


图 5-4 极寒天气年出现天数图

5.4.3 结果分析

首先对 NU 地区年平均温度进行分析，得到如下年平均温度变化图，从图中可以发现，虽然某些年份出现比较大的波动，但是该地区平均温度呈现整体上升的趋势。

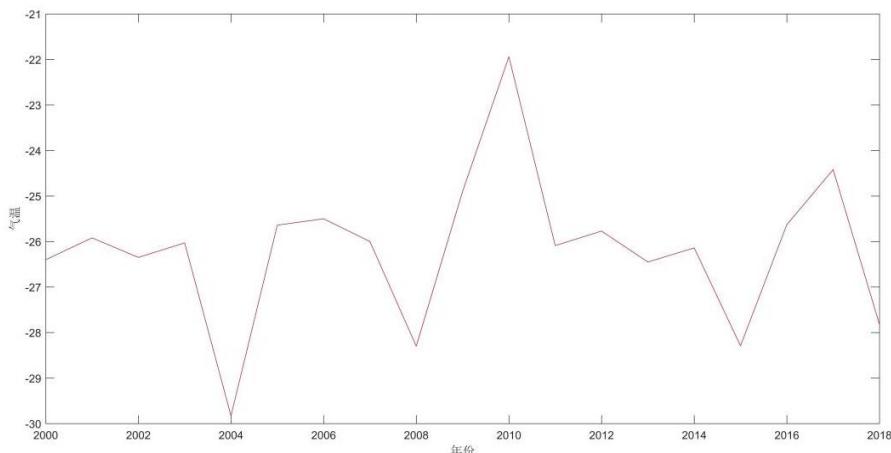


图 5-5 年平均温度变化图

同样对该地区 20 年来年最低温度制作散点图，进行拟合后发现该地区最低温度呈现

上升趋势，也即整体呈现变暖趋势。

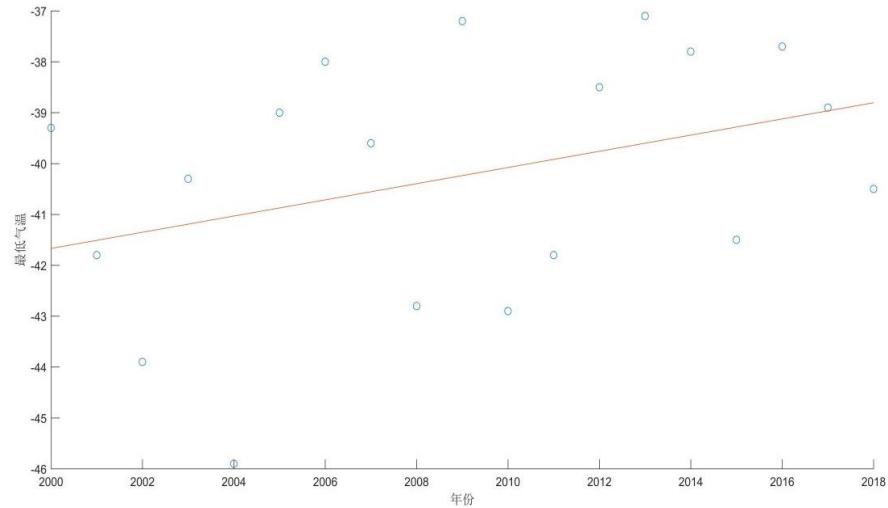


图 5-6 近 20 年最低温度散点图

以上两个图形都能大致说明该地区整体温度呈现上升趋势，但从年最低温度图中可以很清晰的看到极寒天气的出现，近两个十年都出现了 10 次左右，这样就大概说明了该地区整体温度的上升与极寒天气的出现并不矛盾。

应用 Mann-Kendall 方法对该地区极寒天气事件的发生频次的突变性进行检验，M-K 检验结果如下图：

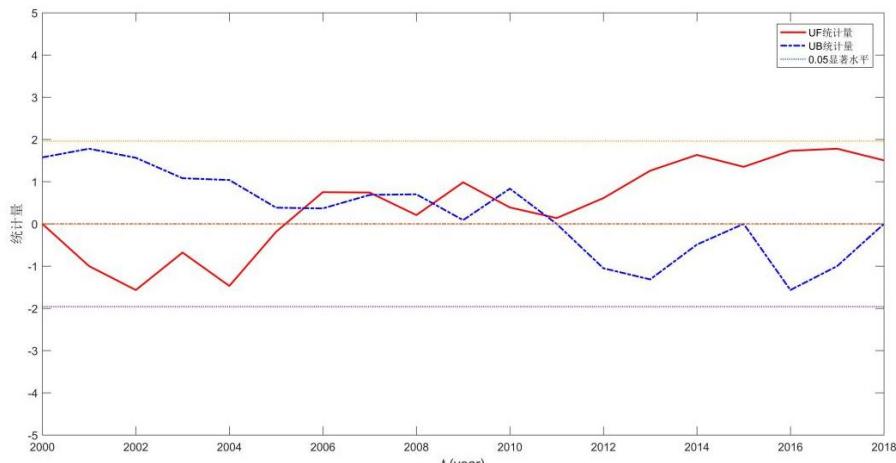


图 5-7 M-K 检验结果图

从图中可以看出，UF 曲线对应的值整体呈上升的趋势，从而说明该地区极寒天气事件的发生频次有上升趋势。但在图上也可以清晰的看到，2005-2010 年前后发生了突变，这也与之前的直方图呈现的结果相互印证。

由于只选取了该地区 20 年左右的数据，加上极寒天气本身发生的样本较少，所以上述分析存在比较大的局限性，虽然可以看出整体上的部分趋势，但这是远远不够的，所以接下来将从其他方面或者角度来继续描述和讨论整体气候变化与极寒天气之间的关系。

5.5 矛盾性探究：分层贝叶斯时空模型

5.5.1 相关理论

(1) 贝叶斯分层模型

一般的，我们把联合分布分解为边缘分布与条件分布的积，即

$$P(AB) = P(A|B)P(B) = P(B|A)P(A) \quad (5-17)$$

从而有贝叶斯定理

$$P(AB) = \frac{P(B|A)P(A)}{P(B)} \quad (5-18)$$

在给定事件 B 的情况下， $P(A|B)$ 的形状不受到因子 $P(B)$ 的影响，条件分布与联合分布仅仅相差一个常数。

贝叶斯推断理论将参数看作变量，依据经验给出先验分布后辅助样本信息得出后验密度，即

$$\pi(\theta|y) = \frac{f(y|\theta)\pi(\theta)}{\int f(y|\theta)\pi(\theta)d\theta} \quad (5-19)$$

其中， $f(y|\theta)$ 为 y 的条件分布函数， $\pi(\theta)$ 为先验密度， $\pi(\theta|y)$ 为 θ 的后验密度。

在分层贝叶斯模型中，我们涉及到超先验的参数，即 $\pi(\theta)$ 的先验分布分两步给出，其中 $\pi(\theta)$ 的先验分布中包含另一个参数 α ，即先验分布为 $\pi(\theta|\alpha)$ ，关于参数 α 我们可以给出超先验分布 $h(\alpha|y)$ ，其中超参数 y 的分布给定。

根据贝叶斯理论，我们可以写出分层情况下的后验分布，

$$\pi(\theta|y) = \int \pi(\theta|\alpha, y)\pi(\alpha|y)d\alpha \quad (5-20)$$

其中，

$$\begin{aligned} \pi(\theta|\alpha, y) &= \frac{f(y|\theta)\pi(\theta|\alpha)}{m(y|\alpha)} \\ \pi(\alpha|y) &= \frac{m(y|\alpha)h(\alpha)}{\int \int f(y|\theta)\pi(\theta|\alpha)h(\alpha)d\alpha d\theta} \end{aligned} \quad (5-21)$$

$$\begin{aligned} m(y|\alpha) &= \int \int f(y|\theta)\pi(\theta|\alpha)d\theta \\ \pi(\theta|y) &= \frac{f(y|\theta)\int \pi(\theta|\alpha)h(\alpha)d\alpha}{\int \int f(y|\theta)\pi(\theta|\alpha)h(\alpha)d\alpha d\theta} \end{aligned} \quad (5-22)$$

在实际操作中，后验分布往往十分复杂，不具有标准形式难以直接计算，MCMC (Markov Chain Monte Carlo) 方法是解决这一问题行之有效的方法之一，其基本思想是通过建立一个平稳分布为 $\pi(x)$ 的 Markov 链来得到一个样本，基于这个样本进行各种统计推断。本题选用 MCMC 方法中的 Gibbs 抽样，具体理论如下：

设 $X = (X_1, \dots, X_n)$ 的密度函数为 $\pi(x)$ ，对任意指定的 $T \subset N$ ，记

$$x_T = \{x_i, i \in T\}, x_{-T} = \{x_i, i \notin T\} \quad (5-23)$$

称 $\pi(x_T | x_{-T})$ 为满条件分布，在指定 $x_T = x_{-T}$ 的条件下，定义如下随机变量：

$$X' = (X'_1, \dots, X'_n) : X'_{-T} = X_{-T} \quad (5-24)$$

而 X'_T 具有密度函数 $\pi(x'_T | x_{-T})$ ，则对任意可测集 B ，

$$P(X' \in B) = \int_{\pi} x'_{-T} \pi(x'_T | x_{-T}) dx' = \int_B \pi(x') dx' = \pi(B) \quad (5-25)$$

由此， X' 的密度函数与 X 相同，均为 $\pi(x)$ 。在这个过程中我们定义了一个从 X 到 X' 的转移核，且其相应的恶平稳分布为 π 。以上的构造方法为 Gibbs 抽样。

(2) 贝叶斯时空模型

基本的思路是将复杂的统计建模问题简化为一步一步的层次模型。基本的分层模型分为以下两个阶段：

- ① 数据模型 [数据 | 过程, 参数] ;
- ② 过程模型 [过程 | 参数] 。

第一阶段我们关注的是数据的观测过程和确定数据分布的模型以及其相关参数，第二阶段关注在第一阶段条件下的其他参数，比如时间效应、空间效应、其他影响因素等，将这些变量与第一阶段的参数连接起来。由于我们的变量是极寒天气的频数，所以可选用泊松分布进行描述。

$$Y_{ij} \sim Poisson(\lambda_{ij}) \quad (5-26)$$

其中， $i=1,2,\dots,m$ 表示空间维度即选取的出现过极寒天气的第 i 个地区（以国家为单位）， $j=1,2,\dots,n$ 表示时间维度即第 j 年。 Y_{ij} 表示第 i 个地区第 j 年发生的极寒天气频次，假设其服从参数为 λ_{ij} 的泊松分布。

第二阶段，建立时间、空间、以及时空交互作用与泊松强度间的关系：

$$\log \lambda_{ij} = \alpha_0 + a_i + b_i + t_j + \text{int}_{ij} + \beta X_{ij} \quad (5-27)$$

其中 α_0 为常数项， b_i 表示空间相关异质效应， a_i 表示不相关空间异质效应即表示的是除去空间分布外，剩下的误差来源， t_j 为时间的自回归效应， int_{ij} 表示时空交互效应， X_{ij} 为协变量。

5.5.2 数据来源和处理

本节所用数据来源于 <http://berkeleyearth.org/data/>，从该网站得到按国家的全球陆地平均气温数据集 Global Land Temperatures By Country. csv 以及全球陆地和海洋气温数据 Global Temperatures. csv。通过筛选和整理得到近 10 年出现极寒天气地区的极寒天气频数信息。

5.5.3 模型建立

在本题之中，协变量部分选择地区年平均温度作为标志全球气候变化的指标，来考虑其对于极寒天气出现频次的影响。

包含时空效应与协变量部分的具体模型如下：

$$Y_{ij} \sim Poisson(\lambda_{ij}) \quad (5-28)$$

$$\log \lambda_{ij} = \alpha_0 + a_i + b_i + t_j + \text{int}_{ij} + \beta Temp_{ij} \quad (5-29)$$

其中， $i=1,2,\dots,m$ 表示空间维度即选取的出现过极寒天气的第 i 个地区， $j=1,2,\dots,n$ 表示时间维度即第 j 年。 $Temp_{ij}$ 表示第 i 个地区、第 j 年的年平均温度。

先验分布假设：

定义 b_i 服从条件自回归，即：

$$p(b_i | b_l, l \neq i) = N(\bar{b}_i, \sigma_b^2 / n_i) \quad (5-30)$$

其中， $\bar{b}_i = \sum_{s \in \gamma_i} b_s / n_i$ ， γ_i 是地区*i*所有相邻区域集合， n_i 为集合中的元素个数。这个条件自回归的随机变量描述了空间上的影响，即更相邻区域具有更接近的极寒天气频数。

a_i 则表示其他误差来源，其先验分布假设为零均值的正态分布：

$$a_i \sim N(0, \sigma_{a_i}^2) \quad (5-31)$$

其中， σ_{a_i} 的超先验分布均假设为均匀分布。

t_j 为的AR(1)的时间自回归效应：

$$t_j | t_{j-1} \sim N(t_{j-1}, \sigma_{t_j}^2) \quad (5-32)$$

其中， σ_{t_j} 的超先验分布均假设为均匀分布。

交互项 int_{ij} 的先验分布假设为正态分布：

$$\text{int}_{ij} \sim N(0, \sigma_{\text{int}_i}^2). \quad (5-33)$$

其中， $i=1, 2, \dots, m$ 表示空间维度，

$$\text{int}_{ij} \sim N(\text{int}_{i,j-1}, \sigma_{\text{int}_j}^2) \quad (5-34)$$

其中， $j=1, 2, \dots, n$ 表示时间维度，即交互项为受到时空因素共同影响的变量。 $\sigma_{\text{int}_i}, \sigma_{\text{int}_j}$ 的超先验分布假设为均匀分布。

协变量系数 β 的先验分布假设为正态分布：

$$\beta \sim N(\beta_0, \sigma_\beta^2). \quad (5-35)$$

其中 β_0 为初值常数， σ_β 的超先验分布假设为均匀分布。

截距项 α_0 的先验分布假设为无信息先验分布。方差的超先验分布假设为均匀分布。超先验分布中均匀分布选择 $u(1, 10)$ 。

在本题之中，我们选用 Gibbs 抽样对参数进行估计，在得到各参数的联合后验密度后，通过抽样将高维后验密度分解为低维。待估参数为 $x = (x_1, \dots, x_m)$ ，共*m*个，其中参数 x_1 的分布依赖于超参数 x_2, \dots, x_m 。使用该抽样方法，需首先给出各参数初始值 $x^{(0)} = (x_1^{(0)}, \dots, x_m^{(0)})$ ，在该抽样的第*n*次迭代中，我们从参数 x_1 开始抽样，其满条件分布形式如下：

$$p(x_1^{(n)} | x_2^{(n-1)}, \dots, x_m^{(n-1)}, y) \propto p(y | x_1^{(n-1)}) p(x_1^{(n-1)} | x_2^{(n-1)}, \dots, x_m^{(n-1)}). \quad (5-36)$$

其中上标为*n-1*的参数值取当前值，上标为*n*的参数值为更新值，即从该分布为参数 x_1 抽取了一个值。

接下来对参数 x_2, \dots, x_m 从其满条件分布中进行抽样：

$$\begin{aligned} p(x_2^{(n)} | x_1^{(n)}, x_3^{(n-1)}, \dots, x_m^{(n-1)}, y) &\propto p(x_1^{(n)} | x_2^{(n-1)}, \dots, x_m^{(n-1)}) p(x_2^{(n-1)}), \\ &\dots, \\ p(x_m^{(n)} | x_1^{(n)}, \dots, x_{m-1}^{(n)}, y) &\propto p(x_1^{(n)} | x_2^{(n)}, \dots, x_{m-1}^{(n)}, x_m^{(n-1)}) p(x_m^{(n-1)}). \end{aligned} \quad (5-37)$$

我们使用上一步得到的 x_1 的更新值，完成了对 x_2 的更新，对后续参数继续进行，直到对所有参数完成更新，进入 Gibbs 抽样的下一次迭代。

5.5.4 模型求解及结果分析

通过反复迭代，可以得到参数 λ_i 的一些自相关函数图以及核密度图。从图中可以看出，当迭代次数达到一定数量后，参数的迭代基本稳定，自相关函数值趋近于零，核密度估计图近似正态，由此我们可以认为迭代已经收敛。

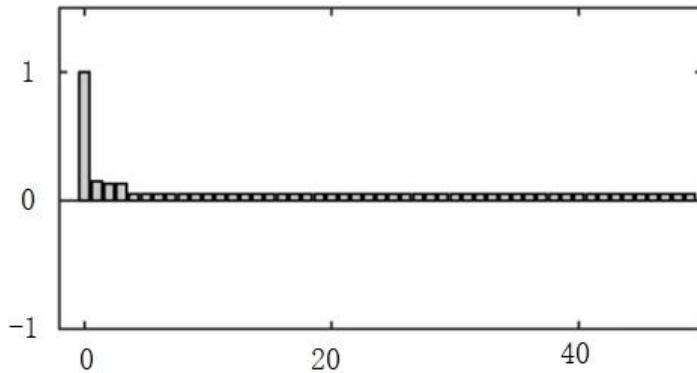


图 5-8 自相关函数图

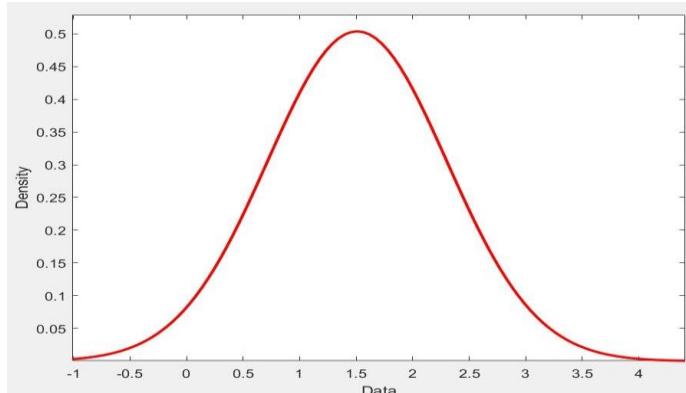


图 5-9 核密度图

此外，我们通过 MCMC 方法可以恰当的估计出温度参数 $Temp$ 的系数值 β 为正，并通过马尔科夫链的收敛性来保证估计的有效性，这说明地区年平均气温对极寒天气频次有着正向的作用，也就是全球气候变暖影响了局部极寒天气事件的发生，全球气候变暖使得极寒天气事件增加。

5.6 本章小结

进入 21 世纪以来，10 年间全球全年平均气温变化非常小，出现了全球变暖停滞现象，这让公众对全球变暖产生了怀疑。而全球多地出现的极寒天气事件，成为人们怀疑全球变暖的依据之一。公众可能会认为，全球变暖和局部极寒天气的发生是矛盾的。也就是说如果全球变暖，那么极寒天气事件不应该出现，因此现在极寒天气的出现只能归结为全球变暖的停滞。如果能说明全球变暖会导致局部极寒天气事件的发生，那么就能说明全球变暖与局部极寒之间不矛盾。

本文从多个角度出发，分析全球变暖对极寒天气发生的影响。

(1) 从地理特征的角度，分析极寒天气出现的基本要素，并探讨了其与全球的气候变化的关系。

(2) 运用极值理论，以加拿大 NUNAVUT 地区为例，分析该地区近 20 年来极寒天气的出现与该地区气候变化的关系，结果初步表明该地区气候变暖对极寒天气有影响。但是以一个地区来分析样本点太少，只能大概的得出之间的关联，具有很大的巧合性和不确定

性。

(3) 最后，通过运用构建分层贝叶斯时空模型，克服极值理论运用中的问题，来分析年平均气温变化对极寒天气出现频次的影响。结果表明年平均气温对极寒天气频次有着正向的作用，也就是全球气候变暖影响了局部极寒天气事件的发生，全球气候变暖使得极寒天气事件增加。

综上所述，全球变暖和局地极寒现象的出现存在关系，并且相互之间并不矛盾。

6. 问题四：“全球变暖”——沮丧？彷徨？希望？

怎么理解“全球变暖了，某地今年的冬天特别冷”？首先，这个说法并不矛盾，这是因为观察的角度和范围不一样。全球变暖是长时间的平均气候状态，在短时间内变化不大，人们一般感受不到。而某地今年冬天特别冷是短时段内的状态，人们可以直接感受得到。其次，“全球变暖”与“某地今年冬天特别冷”联系紧密，全球变暖可能会导致某地今年的冬天特别冷，是其形成的原因之一。

如何探究“全球变暖”新的概念，反映气候变化的趋势和复杂性呢？让我们从系统动力学的角度对气候变化的现象逐一分析。

6.1 全球气候系统概况

气候的变化使得地球创造着干湿冷暖多彩的景象，四季的更替使得人们享受着春夏秋冬不同的韵味。那么，气候变化的趋势到底是什么样的呢？又拥有着什么样的复杂性呢？首先我们应当回到最初的问题，了解气候系统的构成，它是一个完整、相互关联的具有复杂代谢和自身调节机制的系统，包括复杂的物理、化学、生物相互作用过程，从而构成复杂的地球生命支持系统。我们了解到一般所指的全球气候系统应当包括五个基本分系统，分别是大气圈、水圈、冰冻圈、岩石圈（又称陆圈）以及生物圈，随着时间的演变，自身的动力学作用以及外部作用的共同推动之下，呈现出不同时空尺度的气候变化与变化率。

6.1.1 气候系统的构成

大气圈（atmosphere）是因重力关系而围绕着地球的一层混合气体，是地球最外部的气体圈层，包围着海洋和陆地，大气层的厚度大约在 1000 千米以上，但没有明显的界限。它是整个气候系统中最不稳定的存在，而且变化最快最迅速。但是大气圈不但受到其他四个圈层的直接作用和影响，而且与人类活动有着密切的关系。人类主要生活在大气圈中，大气圈的状态和变化直接影响着人类的生存条件和各种活动，而且气候系统中其他圈层变化产生的最后影响结果都会反映在大气圈中。所以，我们认为大气圈是气候系统的中心。

对于大气的成分，主要是由各种气体、水汽，以及固、液体质点（气溶胶）、云等组成。像氮、氧、氩等这些所谓的惰性气体，一般来说，它们与入射的太阳辐射相互作用甚小，与地球发射的红外长波辐射无相互作用，既不吸收也不发射热辐射。但是二氧化碳、甲烷、氧化亚氮和臭氧这类气体，虽然只占大气总体积混合比的 0.1% 以下，但是它们吸收和发射辐射，在地球能量收支中起着基本的作用。不可忽略的是，水汽也是一种自然的温室气体，它可以通过相变转化成水滴、云滴与冰晶，因而对地球气候的变化影响很大。其他的话，主要像大气溶胶，可以把一部分太阳辐射散射向太空而引起地面冷却，但如果是黑碳气溶胶，则它们又可吸收太阳短波辐射而使地表与大气增暖，这是直接作用；另外，气溶胶也可作为种凝结核而形成云滴，间接作用表征辐射收支的变化对气候状态的影响。

大气环流是指大范围的大气运动，其持续的异常状态往往会影响着气候变化。大气系统通过吸收太阳短波辐射以及沿地面向上的感热和潜热输送而获得能量；同时向宇宙空间

放射长波辐射而失去能量。但是由于地球的形状（曲率）和旋转特征，加上大气中主要吸收气体的分布特征，大气的净能量收支存在明显的纬度变化，在低纬度地区有能量的盈余，在高纬变地区有能量的亏损。为了维持大气的能量平衡，就需要有从低纬度向极地的能量输送，还减小了因辐射过程造成的经向温度梯度。

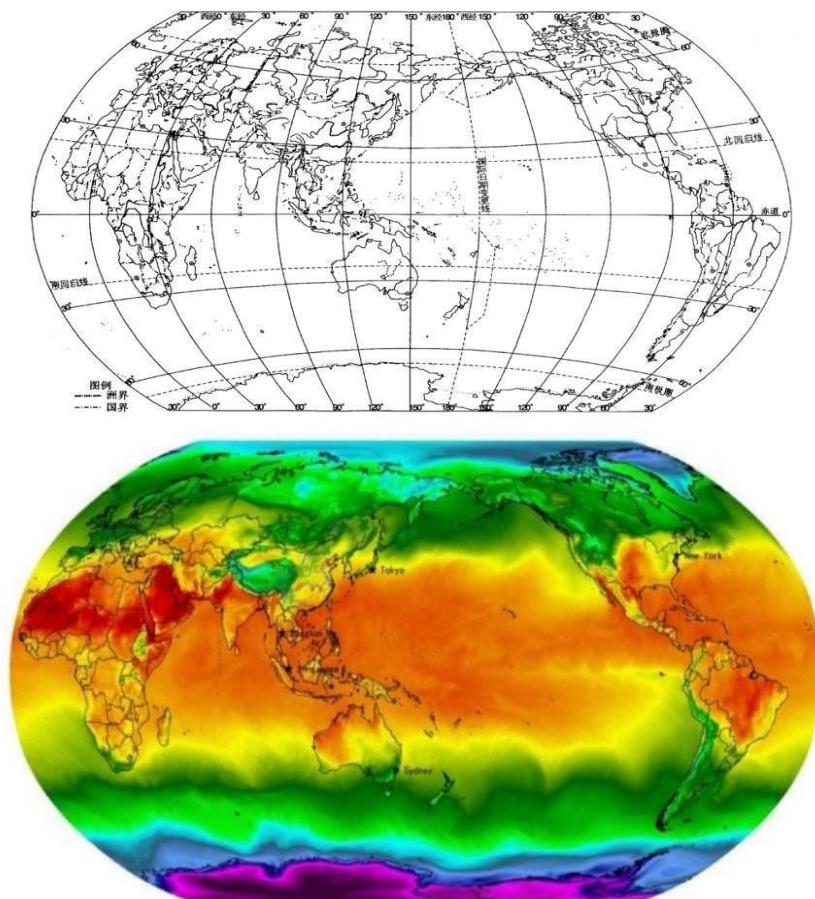


图 6-1 全球温度在经纬度方向上的变化特征

水圈（hydrosphere）是指由地球表面上下，液态、气体和固态的水形成一个几乎连续的、但不规则的圈层。上界可达大气对流层顶部，下界至深层地下水的下限，包括大气中的水汽、地表水、土壤水、地下水和生物体内的水。水圈中大部分水以液态形式储存于海洋、河流、湖泊、水库、沼泽及土壤中；部分水以固态形式存在于极地的广大冰原、冰川、积雪和冻土中；水汽主要存在于大气中，三者常通过热量交换而部分相互转化。

水圈的主体是世界大洋，面积占全球面积的约 71%，因此在水圈中，其对气候变化的影响最大。从大气运动及其变化的角度，海洋具有以下的显著特征：

- (1) 全球海洋吸收的太阳辐射量约占进入地球总辐射量的 70% 左右。
- (2) 海洋有着极大的热容量，相对于大气运动而言，海洋运动比较稳定，运动和变化相对来说比较缓慢，那是因为它是由盐分与温度梯度产生的密度差（即热盐环流）驱动的。
- (3) 海洋可以溶解和储存大量的二氧化碳，是全球碳循环中非常重要的部分。最新的数据表明，海洋每年可以吸收 17 亿吨碳，占矿物燃料燃烧和工业生产排放总量的 27%。对于研究海洋对气候变化的影响，主要考虑的方面总结起来由以下几大类：
- (4) 对系统热平衡的影响。海洋吸收的大量太阳入射辐射，绝大部分（相关数据表示达到 85%）被贮存在海洋表层。这些被贮存的能量会以潜热、长波辐射和感热交换的形式输送给大气，驱动大气的运动。因此，海洋热状况的变化以及海面蒸发的强度如何都将

对大气运动的能量发生重要影响，从而引起气候的变化。

除此之外，各种类型的海洋环流在系统能量输送和平衡中起着重要的作用。在地球上低纬地区获得的净辐射能多于高纬地区，因此，要保持能量平衡，必须有能量从低纬地区向高纬地区输送。据有关资料显示，通过卫星观测，全球有超过 30% 的经向能量输送由海洋承担，其他的主要是由大气完成。

(1) 对水汽循环的影响。大气中 86% 的水汽量由海洋供给，低纬度的海洋是大气中水汽的主要来源地。因此，不同的海洋状况通过蒸发和凝结过程将对气候产生不同程度的影响。

(2) 对大气运动的调谐。由于海洋的运动和一系列的变化具有明显的缓慢性和持续性，相当于本身是一个“储存盘”，具有较强的记忆性，可以把大气环流的变化通过海气相互作用把信息储存起来，然后对大气运动产生作用。另一方面，海洋的热惯性使得海洋状况的变化相比较于陆地具有滞后现象；同时海气耦合作用还可以使较高频率的大气变化减频，耦合波的频率变低后再作用于大气，总的来说就是相当于大气中的较高频变化放缓。

(3) 对温室效应的缓解。尤其是海洋环流，不仅减小了低纬大气的增热，使高纬大气加热，降水量亦发生相应的改变，而且由于海洋环流对热量的向极输送所引起的大气环流的变化，还使得大气对某些因素变化的敏感性降低，大气中二氧化碳含量增加产生的温室效应就是因为海洋的存在而被减弱。

最能体现海洋与气候变化相关的两大特征是海面温度 (SST) 和大洋环流，通过问题一的分析，我们发现 SST 数据与大气环流和气候异常的关系十分密切，最典型的就有厄尔尼诺现象和拉尼娜现象。进一步分析大洋环流的特征，首先大洋环流引起的海洋热量的向极输送，低纬度海洋获得热量而在高纬度海洋失去热量，同时低纬度大气被冷却，高纬度大气被加热。热量平衡的分析表明，冰雪反馈作用使得高纬度大气增暖，但稳定的大气层结又阻止了热量向对流层高层的传送。因此，不仅对于海洋状况而且对于大气环流和气候也有明显的影响。

冰冻圈 (cryosphere) 指地球表层连续分布并具有一定厚度的负温圈层，包括冰川（含冰盖和冰帽）、河冰、湖冰、积雪、冰架、冰山、海冰，以及多年冻土和季节冻土（表）。大气圈内的雪花、冰晶、冰雹、霰等固态水也是冰冻圈的组成部分。冰冻圈在赤道附近海拔最高，向南北两极逐渐降低到海平面。目前，冰川覆盖了全球地表约 3% 的面积，储存了 75% 的非海洋水(淡水)。海冰面积占冰冻圈面积的 7%，永久冻土占陆面面积的 20%~25%。

冰冻圈对气候系统之所以重要是由于它对太阳辐射有较高的反射率(反照率)、低的热传导率、大的热惯性，以及在驱动深海环流中的关键作用。它能影响地表能量与水汽通量、云、降水、水文循环以及大气与海洋环流，但最明显的影响是对海平面高度的影响。因为冰原储存了大量的水，其体积变化可引起海平面上升。如果包含近 90% 世界冰川的南极冰原全部融化，那么全球海平面将可能升高 70m。

冰雪覆盖的变化也是气候学家用来制作不同时间尺度气候预测的重要因子，冰冻圈对大气环流的影响主要表现在其对主要气候型或大气活动中心的影响上。由季节冰雪覆盖产生的能量与水循环的滞后作用、存储于冰架和冰川中的水体，以及封存在冻土中的温室气体也是冰冻圈对气候产生作用的重要因素。正是通过这些因子及相关的反馈过程，冰冻圈在全球气候及气候模式中起着重要作用。

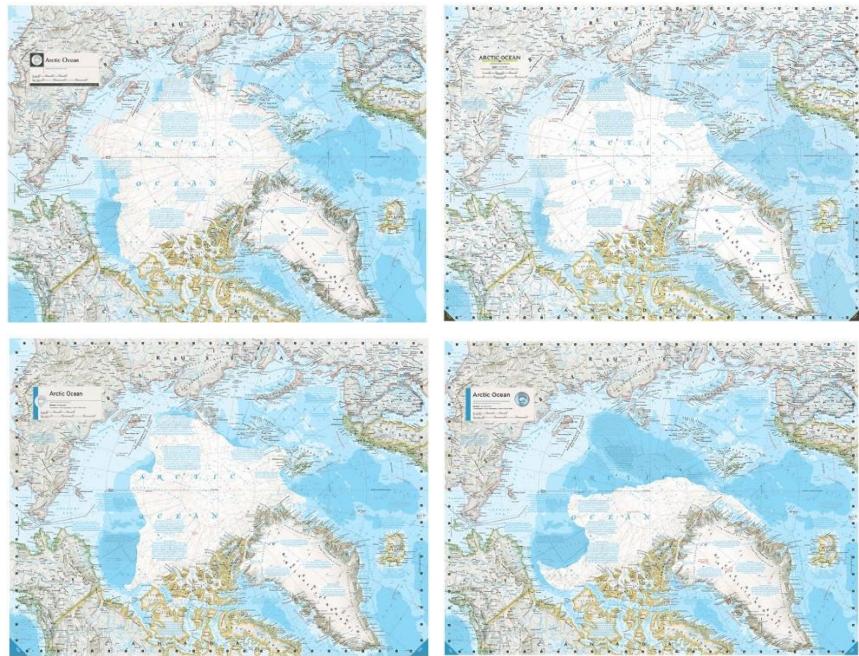


图 6-2 1979-2016 年北极冰雪覆盖面积的变化

积雪异常也具有明显的气候效应，除了因其大的反照率造成积雪区与无雪区的净辐射差很大，地温和气温也相差很大，从而影响基本风系的强度和位置之外，积雪融化后将使土壤湿度增大，蒸发量也会加大，大气中的热量和水分平衡也会发生改变。而且积雪融化及其影响有一定的持续性，可持续数月之久。

岩石圈（lithosphere）是指固体地球的上层部分。火山活动虽然是岩石圈的一部分，但不包含在气候系统之中，而是作为一种自然的外界强迫因子影响地球的气候。与气候变化最密切相关的部分是陆面的植被与土壤以及相关联的陆面过程，它们能够影响从太阳接收到的能量中又有多少返回到大气中，其中有些是以长波辐射的形式回到大气中，并随着陆面增暖可使大气加热；有些是通过土壤或植物的叶子蒸发或蒸散水分，因为土壤水分蒸发时需要能量或吸收热量，因而土壤水分或湿度对地表温度有强烈的影响。

陆面的结构或其粗糙度在风吹过陆面的时候也可从动力学上影响大气，地表面的粗糙度会改变地表的风，影响大气边界层的动力特征和地气物质交换。沙尘暴就是一个很好的例子，由于植被的减少，影响了区域大气辐射的收支。不同景观的陆面到底能吸收多少太阳辐射，实际上主要取决于植被的状况。例如沙漠比植被覆盖区反射更多的太阳辐射。在植被区中，草地比森林又可反射更多的辐射。当太阳辐射入射角较低时（如在高纬度的冬季），植被对反照率的影响会增大。当积雪覆盖地面时，反照率也会增大。森林是雪上的一种光吸收层，而裸露的地面和草则不是。

土壤和植被对气候的另一个作用是吸收和产生温室气体，以此影响大气的红外辐射收支。植被可释放能进行化学反应的有机气体，以后通过大气中的反应产生对流层温室气体臭氧。大气低层发生的光化过程也可使植物放出的碳氢化合物形成小颗粒。这些颗粒具有散射光线的特性，能够降低大气层对太阳辐射的透过率。此外，植被也能影响空气中尘埃的量，从而影响光的折射。

植物在一定程度上控制着陆面的水文循环，植物的叶子在光合作用时开启它们的气孔，使其失去水汽，吸入二氧化碳，更有数据表明大约陆地上 1/3 的降水是由植物再循环的水汽供应。但是当植物由于自身生长需要时，气孔关闭减少水分损失，周围环境的空气也会受到影响，产生明显增暖的效果。土壤的重要作用之一是储存水分，并因此影响水分从陆面蒸发的时间，这样的蒸散也影响着热量的传递。

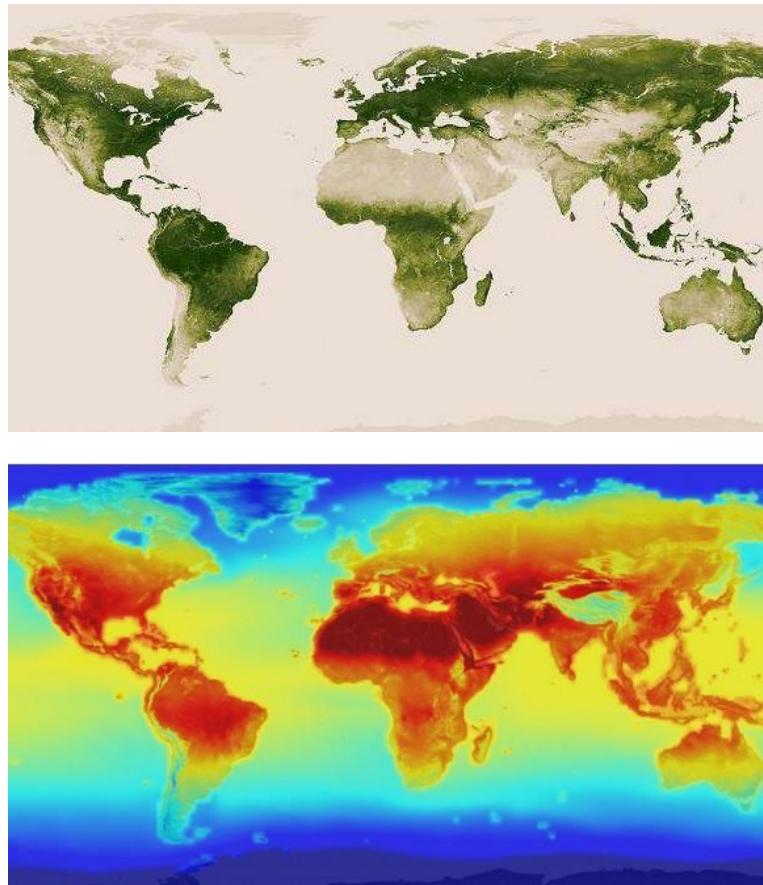


图 6-3 全球植被覆盖情况及其对应的温度变化特征

生物圈（biosphere）包括陆地和海洋以及所有的生态系统和生物。通过生物圈的生物过程与物理和化学过程的强烈相互作用可以产生维持地球上生命系统赖以生存的环境。生物圈对大气的成分有着明显的影响，生物能够直接地转化和释放温室气体。通过光合作用，海洋和陆地植物，特别是森林能够把二氧化碳转化并储存足够量的碳。于是，生物圈在碳循环过程和其他气体的循环过程中扮演了重要的角色。由于气候影响碳的储存和微量气体的交换，这同样会发气候变化与微量气体浓度之间的反馈过程。

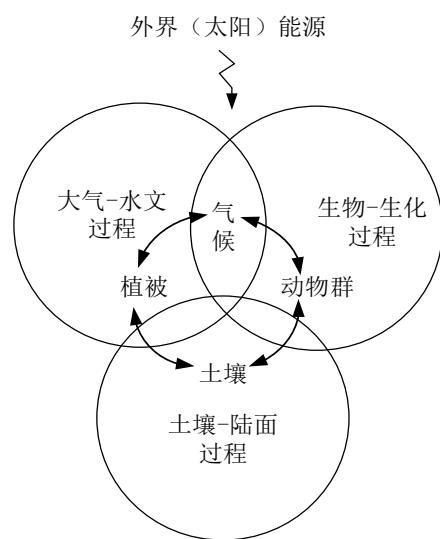


图 6-4 生态系统中物质交换的主要过程

另外，生物对太阳光的反射改变着大气的辐射平衡，同时通过蒸发和蒸腾影响着进入大气的水分。人类作为生物圈的组成部分，其活动对大气成分的影响重大，例如南极附近臭氧分压存在季节性的变化，北半球的冬季南极会出现臭氧洞的现象。并且这些臭氧气压也会发生季节性的变化以及年际的变化。

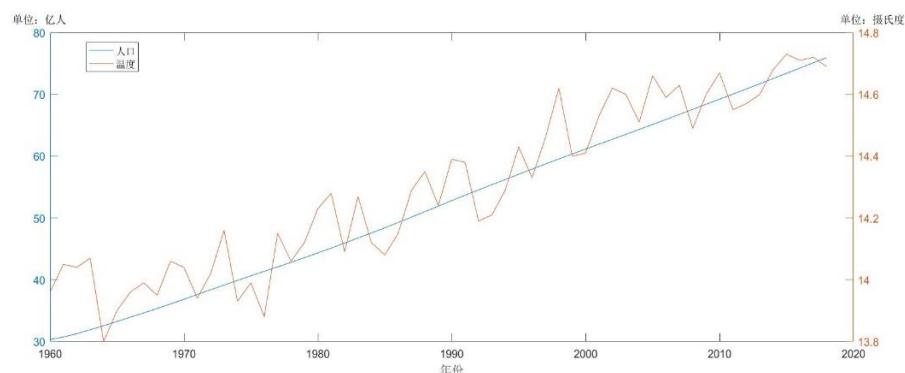


图 6-5 人口增长和全球变暖的关系

6.1.2 气候系统内各圈层的相互作用

气候系统的各个圈层不是独立存在的，它们之间发生着明显的相互作用，这种相互作用不但有物理的、化学的和生物的，还具有不同的时间与空间尺度，从而使得气候系统成为一个复杂的系统。研究长期的气候变化，除讨论气候系统的基本规律及特性之外，各个子系统之间的相互作用正是我们需要关注的重点。其中主要包括：海洋-大气相互作用、陆地-大气相互作用、生物圈-大气相互作用等。而在这些相互作用的过程当中，我们主要考虑的都是能量、动量以及质量交换三个方面对气候变化的影响。

（1）海气相互作用

一直以来，海洋-大气相互作用被认为是地球气候变化问题的核心内容，海洋和大气强烈耦合在一起，并通过感热输送、动量输送和蒸发过程交换热量、水汽和动量。热量和水汽是水文循环的一部分，可产生凝结，形成云、降水与径流，并为天气系统提供运动的能量。另一方面，降水对海洋的盐度及其分布和热盐环流有影响。大气与海洋也交二氧化碳，都是全球碳循环的重要部分。二氧化碳在下沉到深海的极区冷水中溶解，在近赤道较暖的上升海水中释放，从而维持一种平衡。

大气和海洋运动的原动力都来自太阳辐射能，但因为海水反照率比较小，吸收到的太阳短波辐射能较多，而海面上空湿度一般较大，海洋上空的净长波辐射损失又不大。因此，海洋都有比较大的净辐射收入，热带地区海洋面积最大，因此热带海洋在热量贮存方面具有更重要的地位。如图观察一年中到达大气顶的纬向平均能量分布，S1 区域表示陆地区域所占的量，S2 区域表示海洋区域所占的量，显然热带海洋可以得到更多的能量（单位： $10^{22} \text{ J}/5^\circ$ ）。

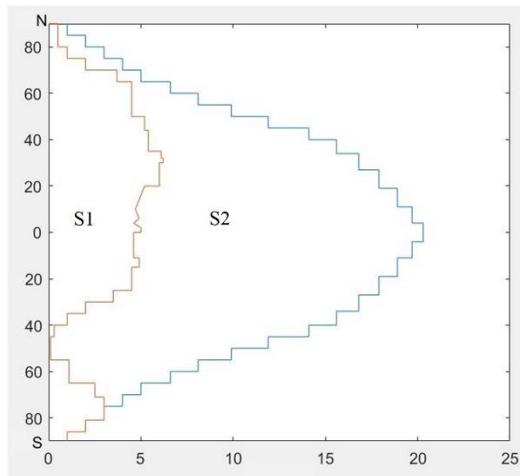


图 6-6 一年中到达地球的纬向平均能量分布

大气对海洋的影响主要是风应力的动力作用，表现出来的典型特征就是：北半球海区，环流沿顺时针方向流动；南半球海区，环流沿逆时针方向流动。还有所谓“西向强化”的特征，最典型的是西北太平洋的近岸海区和北大西洋的西部海域，那里流线密集，流速较大，而大洋的其余部分海区，流线稀疏，流速较小。

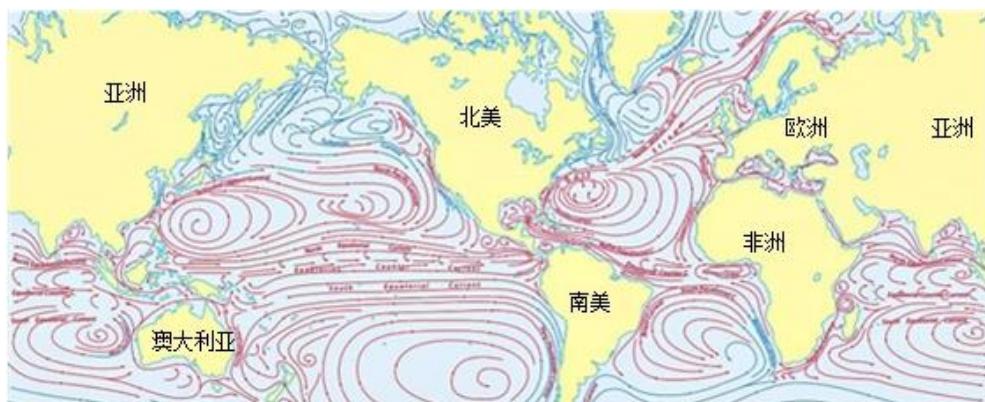


图 6-7 世界大洋环流趋势图

海洋是旋转地球上受重力作用的含盐水体，它的状态和运动既受地球转动和重力的作用，又受到太阳辐射和大气运动的影响。各个海域的海水温度、盐度和密度都有着显著差异，且因太阳辐射的加热作用，海洋上层的温度较高，密度较小，使海水有温度随深度下降而密度随深度增加的特性。但是，其铅直分布却呈现出有规律的层次结构，且一般可分为混合层、跃层和深层，在海水温度、盐度和密度方面分别呈现均匀、跃变和均匀的特性。通过大气和海洋混合层间热量、动量和质量的直接交换，就产生了海气相互作用。对于长期天气和气候的变化问题，都需要知道大气底部边界的情况，尤其是海面温度及海表热量平衡，这就需要知道海洋混合层的情况。

（2）陆气相互作用

陆面过程主要包括地面上的热力过程（辐射及热交换过程）、动量交换过程（摩擦及植被的阻挡等）、水文过程（降水、蒸发和蒸腾、径流等）、地表与大气间的物质交换过程，以及地表以下的热量和水分输送过程。以上诸过程一方面受到大气环流和气候的影响，同时又对大气运动和气候变化有重要的反作用。

生物-地球物理反馈，是一种通过辐射过程而自己加剧其干旱化的自感反馈。比如，以

撒哈拉地区为例，那里本来就少雨，而过度的放牧使地面反照率（30%以上）增大，也就是比四周反射更多的太阳辐射；同时天空少云和地面温度高又造成更多的红外放射。这样，就导致该地区与周围相比是一个辐射汇集区域，为了维持热力平衡，该地区的空气必然被压缩下沉，加剧了干旱，进一步使得植被退化，造成土地沙漠化。

土壤温度和湿度的反馈，这是由于土壤的热容量比空气大得多，土壤的热状况及其变化将对大气的陆面下边界条件起重要作用；同时，土壤湿度会改变地表的蒸发，从而会影响地气间的水分交换以及大气中的潜热释放。具体来说，土壤温度若偏高，后期降水量就偏多，反之亦然；而且较深层的土壤温度所反映的降水量的滞后时间较长。而对于土壤湿度，一般来讲，土壤湿度偏低，则会使地面温度增加，射出长波辐射也增加；同时，比较干的土壤其反射率较大，导致地面吸收的太阳辐射减小。另外，土壤湿度直接联系着蒸发，较潮湿的土壤有利于增加大气的含水量，而且使大气稳定度降低，有利于降水发生。

陆面大部分的地方由各种植物覆盖，不同的植被有其自身的物理和生物特性（反照率、蒸腾特性等），从而使地表过程变得更为复杂。但是概括起来，当降水落到植被表面时，一部分可被植被表面截留，然后被再蒸发到大气中；其余部分滴落到地面后，部分渗入土壤，部分成为径流；渗入土壤的水分还可以有部分渗透到更深层而成为地下水。植物的根可以将土壤中的水吸到茎和叶上，通过蒸腾作用还会有一部分回到大气中。另外，植物冠层的反射和散射作用对大气及地面的辐射过程也有极明显的影响。

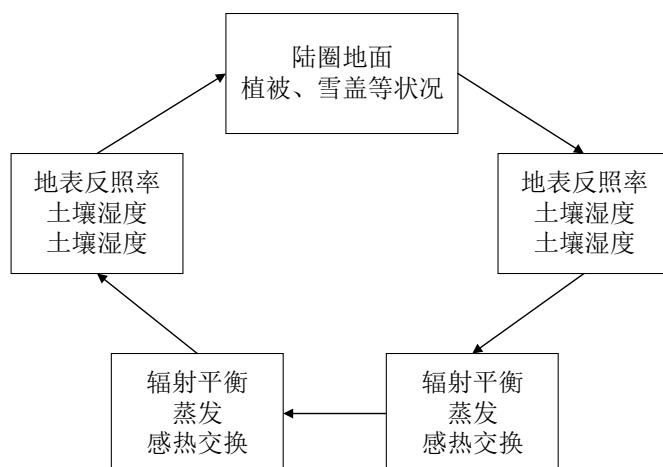


图 6-8 陆圈与大气圈间相互作用示意图

(3) 陆海相互作用

陆地和海洋之间的相互作用最关键的问题是海岸带地区的变化以及跨界能量传输的问题，这其中会涉及到：跨陆海界面的物质输送及沿岸生态系统对气候变化的影响，海岸带的加速变化对来自上游陆地地区的物质转移、过滤或储存的能力的影响，气候系统的系统变化对海岸带特别是最脆弱地区的影响，海气界面对加热场及大气环流的影响等等。

除了上述三种相互作用之外，各圈层间的其他相互作用也值得注意。如海冰可阻碍大气与海洋之间的交换；生物圈通过光合作用与呼吸影响 CO₂ 含量；生物圈通过影响水分向大气的输入；通过改变太阳辐射反射回太空的数量影响大气的辐射平衡。总之相互作用的例子还可以列举出很多，所有这些都说明气候系统包含了非常复杂的物理、化学与生物过程与反馈作用。气候系统中任一圈层的任何变化，不论它是人为的或是自然的，都会通过相互作用造成气候系统的变化或气候的变异。

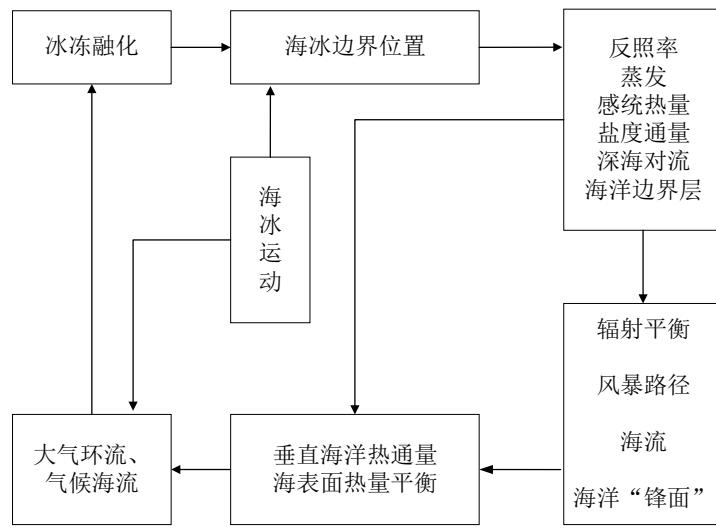


图 6-9 大气圈-冰冻圈-水圈相互作用示意图

6.1.3 气候系统外的影响因素

(1) 太阳活动

大气和海洋运动的原动力是太阳辐射能，大气和海洋获得的太阳辐射能的多少必然对大气和海洋的状况及其运动有重要影响。一般来说我们把到达地球大气外界的直接太阳辐射能总量看作定值；但大气过程也可以造成地球大气系统获得的太阳辐射能有各种变化。太阳活动主要是通过影响大气臭氧层的变化从而影响气候变化的。

(2) 火山爆发

火山活动虽然是岩石圈的一部分，但是不包含在气候系统之内，而是作为一种外界影响因素影响地球的气候。它是地球内部运动的一种表现，然而火山爆发所喷射出的物质可以到达大气，并且可以在大气中存在 1 年或更长时间，从而会改变大的成分以及光学性质。因此，火山爆发也被认为可能是气候变化的重要影响因素之一。

6.1.4 影响气候变化的完整系统

众多因素的相互之间物理、化学、和生物上的相互作用过程，以不同的时空尺度发生在全球气候系统的各个方面，使得这一系统非常复杂。虽然全球气候系统不同部分的特征有所不同，但它们都普遍存在着能量交换的过程，次系统的各个方面都是互有联系的。

分析气候变化考虑的系统，尽可能包括气候系统内部和外部因素的各个方面，系统外部的影响主要包括火山活动和太阳辐射，主要影响大气圈的平流层相关特性。另外，一个主要的因素是人类活动的影响，包括土地利用和温室气体及污染物排放等方面的特征。总的来说，这个完整的系统在各个分量的相互作用之下，形成了不同时空尺度上的气候变化，气候维持一种定常态也是这个系统动态平衡的表现，气候的变化围绕着这个平衡态的浮动。

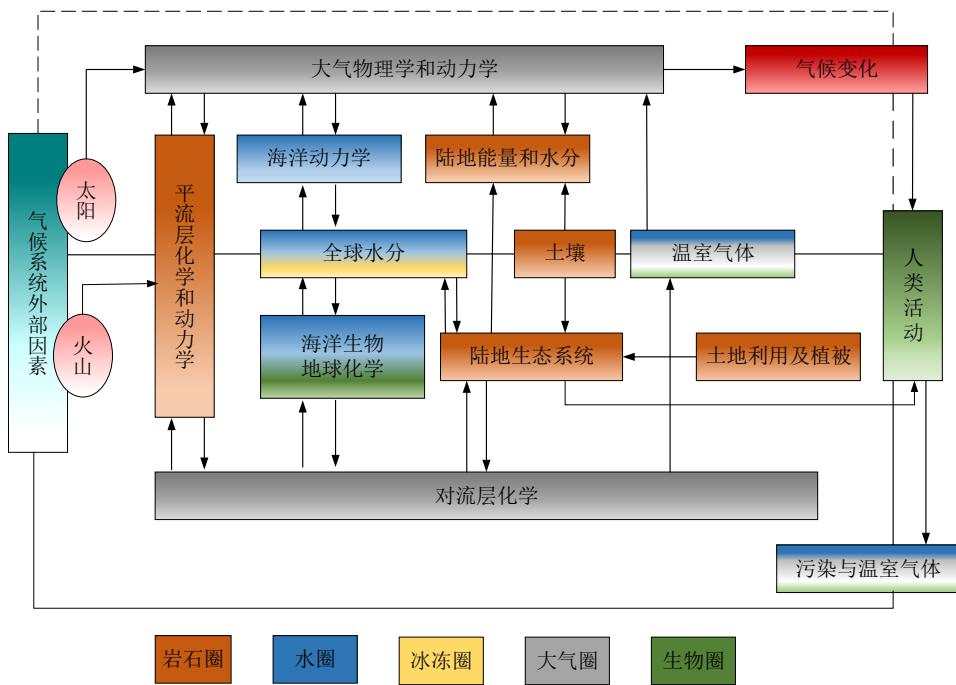


图 6-10 物理气候系统

6.2 从系统动力学角度分析气候变化

6.2.1 系统动力学理论基础

系统动力学（System Dynamics, SD）以分析信息要素之间的因果反馈关系为内容，是一门认识与解决系统、沟通自然科学与社会科学的横向学科。它运用定性与定量相结合、系统综合推理的方法研究复杂系统的问题，通过建立数学模型描述系统内部的结构以及关系，从而处理高阶次、非线性、多重复反馈的复杂时变系统。

系统动力学核心思想认为，系统行为虽复杂多样，在外部环境的作用下千变万化，但其发生发展的模式与特性却主要取决于其内部的动态结构与反馈机制。即内因是系统存在、变化、发展的依据，外因是系统存在、变化、发展的客观条件，系统的演化方向由内、外因通过反馈机制共同决定。系统动力学所研究的目标系统，如本文研究的近三年来恐怖袭击事件系统，是包含反馈环节及其调节作用的系统，要受自身历史行为的影响，把历史行为的后果回授给系统，以影响系统未来的行为。

反馈是系统动力学的核心概念，指系统内同一单元、同一子块内输入与输出间的关系。系统动力学中最基本的结构就是反馈回路，若干相互作用的反馈回路按照一定的系统结构组成一个复杂的系统。因果回路图用于表达系统反馈结构，是系统动力学中的重要工具。它包含因果关系键与因果反馈环。因果关系用一个箭头线表示，自 a 指向 b 的箭头线表示 a 作用于 b, a 称之为原因，b 称之为结果。箭头线称为 ab 间的因果关系键。多个因果关系首尾链接形成反馈回路，反馈回路可分为正反馈回路与负反馈回路，正反馈回路将变化加强，负反馈回路自我校正，抵消扰动。

系统动力学中采用的变量主要有四种，第一种是状态变量，第二种是速率变量，第三种是辅助变量，第四种是常量。状态变量又称水平变量，反映系统物质、信息随时间变化积累过程的变量，其积累量的增加或减少取决于流入与流出的速率。速率变量描述物质流动积累效应变化的快慢，其特征是瞬间性，在模拟中通常采用平均速率替代瞬间速率完成计算。辅助变量介于状态变量和速率变量间，是描述中间变量的变量，其描述的是决策过

程中间环节。常量是系统模拟时间内变化量可忽略不计的量，一般为系统中的标准或局部目标，又或者是通过常量输入流量变量。

运用系统动力学研究现实系统，具体方法步骤如下：

- (1) 对目标系统进行深入地系统思考，进行现实系统抽象与简化，抓住要点；
- (2) 分析目标系统的结构，可能的演化模式，进行因果关系及反馈回路的分析，建立因果回路图，识别状态变量与普通变量；
- (3) 综合运用多种方法，以系统动力学为主，建立系统动力学模型；
- (4) 进行模型的仿真和检验，及各种策略和政策分析；
- (5) 分析模拟结果，获得对了解或改善现实系统有益的结论。

总的来说，这一方法的基本思路是：从问题的系统思考出发，以系统动力学为主结合多种方法，定性分析与定量分析结合，建立综合模型，进行计算机模拟与分析，并进行真实性检验及与实际社会现状的对照，进而获得有意义的结论。

6.2.2 气候变化系统的系统分析

气候变化的表现是一个时段内状态变量相比较前一时段发生了显著的变化，由于气候系统内部以及外部的作用，不同的圈层之间的成分进行相互作用，最终达到一种具有时空特征的现象。气候变化问题本身就是一个涉及多方面因素的多层次、非线性、复杂的系统问题。一般来说，空气中干湿冷暖的变化是气候变化的显著表现，讨论不同圈层之间的作用（即子系统）涉及到的能量交换或许是气候变化的一个突破口。

(1) 明确建模目的

通过建立气候变化演化过程系统动力学模型，分析气候变化演化过程中内部反馈结构及其演化机制，明确各种耦合作用从发生到发展再到结束整个能量交换过程中各因素之间双向影响和非线性关系，再应用仿真计算，研究气候在不同阶段的演化情况，并且在此基础上，一方面要能辨别气候演化过程所处的阶段，另一方面，判定气候发展的主要因素，并对理解气候变化的发生提供客观的依据。

(2) 确定系统边界

从系统动力学角度来看，划分系统边界主要遵循以下两个原则：一是要包含构成系统的主要实体，这些实体在系统内部相互联系，相互作用，构成一个有机的整体；二是描述这些实体的变量相互影响，相互作用，构成反馈机制，形成各种各样的反馈回路或反馈环，系统边界的划分一定要保证这些反馈环的完整性。如果无法做到这一点，则说明系统的边界设定不当，应该扩大边界。

本文为了更好研究气候变化趋势及其复杂性，选取的边界模型为：

- ① 模型主要研究气候变化的演变特征及现象，但不考虑突发性无规律事件（比如行星大爆炸，火星撞地球）。
- ② 各子系统之间相互作用以及能量传递的量化指标根据以往的数据集进行参考，不考虑未来地球内核各种运动产生的影响。
- ③ 除考虑大气系统内各圈层的相互关系，外界因素将火山爆发和太阳活动纳入参考范围，宇宙射线等外界因素不予以考虑。
- ④ 宇宙空间对气候系统的影响主要体现在水、氮气、氧气、二氧化碳、臭氧和气溶胶的含量上，其他的作用我们不考虑。
- ⑤ 我们将地球和太阳本身的运动看作一种规律的演变过程。

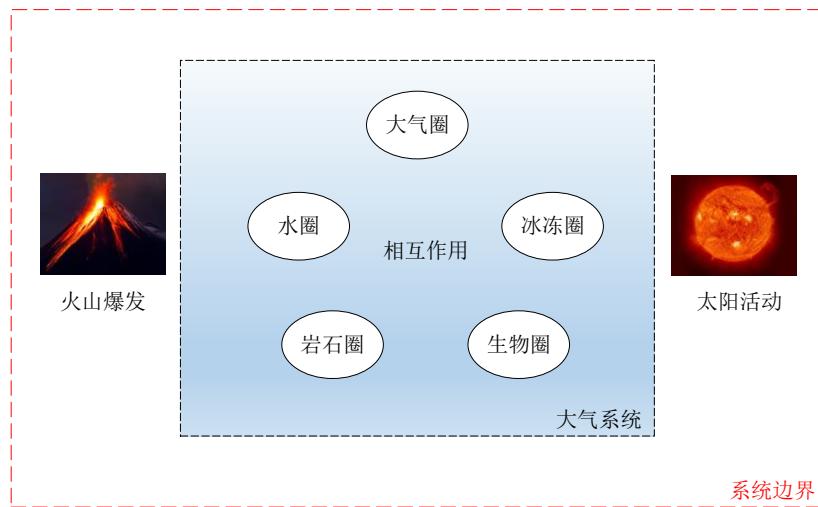


图 6-11 气候变化系统边界图

(3) 选择系统变量

针对气候变化的相关性质，我们可以将该系统进行分阶段分析。首先大气系统内几个单独的圈层对能量输入输出的影响；然后是各圈层之间的耦合作用，对能量交换的影响；最后，不可忽略的外界两个重要的影响因素，包括火山爆发和太阳活动，在能量传递中的作用。每一个子系统之间的影响同样具有反馈作用，不过最后都会体现在大气圈中能量的变化。

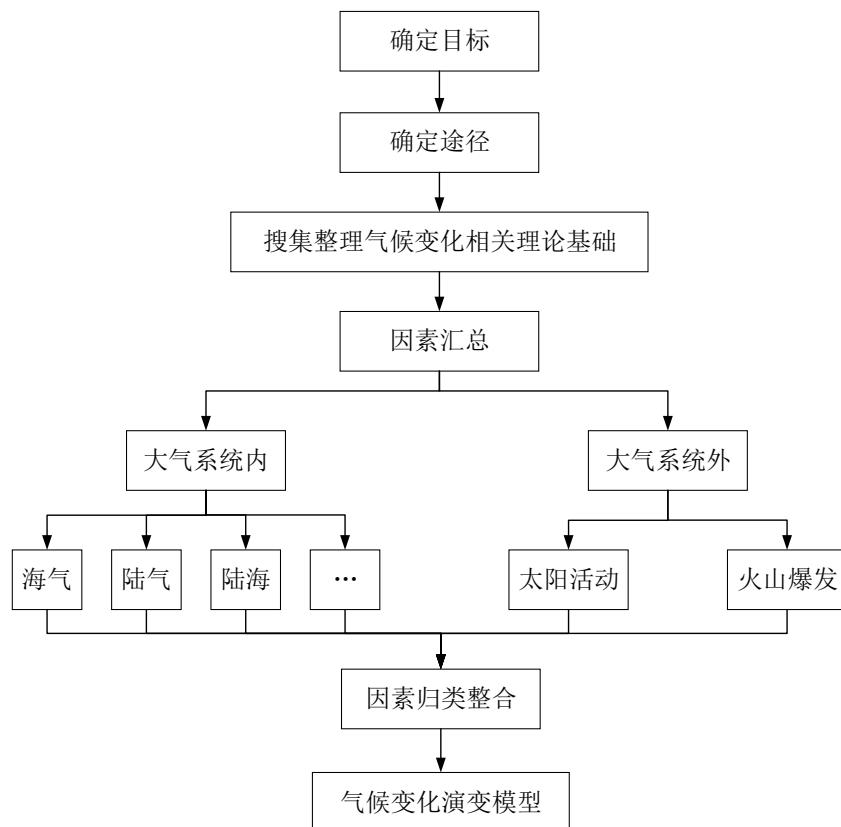


图 6-12 构建气候变化发生演变影响因素的步骤示意图

6.2.3 气候变化 SD 模型构建及分析

气候变化的发生是由于多个子系统的综合影响作用导致的，在建立完整的气候变化模型之前，我们需要明确气候变化演化系统的结构，根据之前分析有关于全球气候变化的基本特征，我们将气候变化分为以下及各子系统，当然这些子系统之间也存在着相互作用的联系。

首先，对于气候变化的特征，有很多参考的因素，我们选取的是温度的变化作为气候变化的考察方式，当然其它例如降水量等也能表示气候的变化。至于分析气候变化的影响因素，回到之前讨论的问题，地球上大气中温度的变化无非就是大气中热能的变化，不同子系统之间物质转化的物理、化学、生物过程都会引起能量的转化，所以考察气候变化的因素是复杂的，涉及到方方面面的影响，对此我们逐个分析各子系统中能量转化的过程。

(1) 考虑有无陆圈作用的海气耦合子系统

海洋和大气的状况在气候尺度上有着密切甚至是共生的关系，目前来说，海气耦合过程仍然是一个复杂的过程，世界上众多的专家和学者并没有对此物理过程有一个直接简单的介绍。但是，从能量转移的角度来说，由于海洋的覆盖面积从太阳透过大气层到达地球表面的辐射大多数都被海洋表面（约占地球总面积的 71%）所接收，因此海洋的热力变化会十分明显，其中极端的现象就是厄尔尼诺现象和拉尼娜现象。

正反馈：对热带海气耦合的不稳定增长，过程表现为赤道东太平洋的一些地区产生赤道中太平洋的西风异常。反过来，这个西风异常产生波浪使得东部海水温度产生巨大的变化。

负反馈：在上述温度的不稳定增长过程中，西风异常也产生赤道影响的另一种波浪，在近赤道外向西传播。当遇到西边界后成为沿赤道上翻的新波浪。这个冷却的波浪到达赤道中太平洋，减小了大气的热通量，迅速使耦合系统进入冷的状态。

在这种模式中，考虑的变量有三个：海水温度异常，西风应力以及温度突变的值。

随着观测对象的增多以及复杂海气耦合模式的发展，多时空尺度下的海气相互作用值得去考虑：

当处于无大陆作用时的全球海气相互作用时，如果有一相对赤道对称的纬向异常西风作用于全球海洋，则风应力通过海洋的效应，赤道纬度带海温升高。海水环流的向极分量形成中纬度及副热带大气西风角动量增大，而向赤道的分量形成低纬度东风的增大。沿赤道的东风发展，导致赤道冷水上翻，海温下降。此时沿赤道的东风减弱，即低纬度异常西风发展，于是又形成赤道附近海温升高。环流的异常变化和赤道内外海温异常之间形成了正负反馈机制。赤道附近海温变化与赤道西风异常有着周期性的差值关系。

经向大陆时的全球海气相互作用会呈现不一样的特征。如果全球海洋被经向的大陆带划分成几个洋，大陆上没有山脉地形。在这样一个海气耦合系统中，若有一相对赤道对称的纬向异常西风作用于全球海洋，则一方面沿赤道附近海温升高，另一方面这几个划分的洋东部海温也升高，属于正反馈特征，而西部海温降低，属于负反馈特征。

(2) 考虑生物圈作用的陆气耦合子系统

陆圈和大气圈之间的能量交换过程主要有辐射及热量交换过程、水循环引起的能量交换过程、动量引起的能量交换过程。正反馈机制包括太阳辐射通过大气层到达陆地表面，陆地表面作为大气圈的下界，承载着较多的能量。并且在生物圈的作用之下，植被通常可以通过光合作用储存能量，生长过程与陆圈交换能量，然后通过呼吸作用将能量与大气圈进行交换，这本身也是一个包括正负反馈的过程。

(3) 考虑生物圈作用的陆海耦合子系统

在海气耦合子系统中，本身就考虑了陆圈的情况。但是由于陆圈和海洋存在明显的交界处，众所周知，沿海气候和内陆气候有着明显的差异，这其中主要就是陆海耦合子系统的作用。典型的特征有，沿海地区的气温往往变化幅度会比内陆的小，生态系统较为脆弱造成对气候变化的极度敏感。此外，高温热浪、强降水以及干旱等极端气候事件的变化对沿海地区的植被覆盖也有着重要的影响，存在明显的正负反馈机制。

(4) 考虑外界影响因素的大气系统

太阳活动对大气系统的影响主要是辐射的变化以及对地球上水循环的影响，正反馈环节主要是太阳辐射的进入使得地球的大气系统的能量显著增加，但是地球辐射也会将一部分能量传输出去。在建模过程中由于属于边界外的区域我们不考虑，所以太阳对整个大气系统是明显的正反馈作用。

火山喷发是地壳运动的一种形式，属于地球内部运动在地表的一种表现形式。火山喷发能够改变附近地区的植被覆盖情况、陆圈的结构、大气圈的气体成分等，如果实在海洋附近，甚至会短暂性改变海洋部分地区的温度特征，期间也存在着明显的正反馈作用。

因此，我们将各个子系统的因果关系分析总结，得到完整的气候变化演化过程总流程图如下所示。

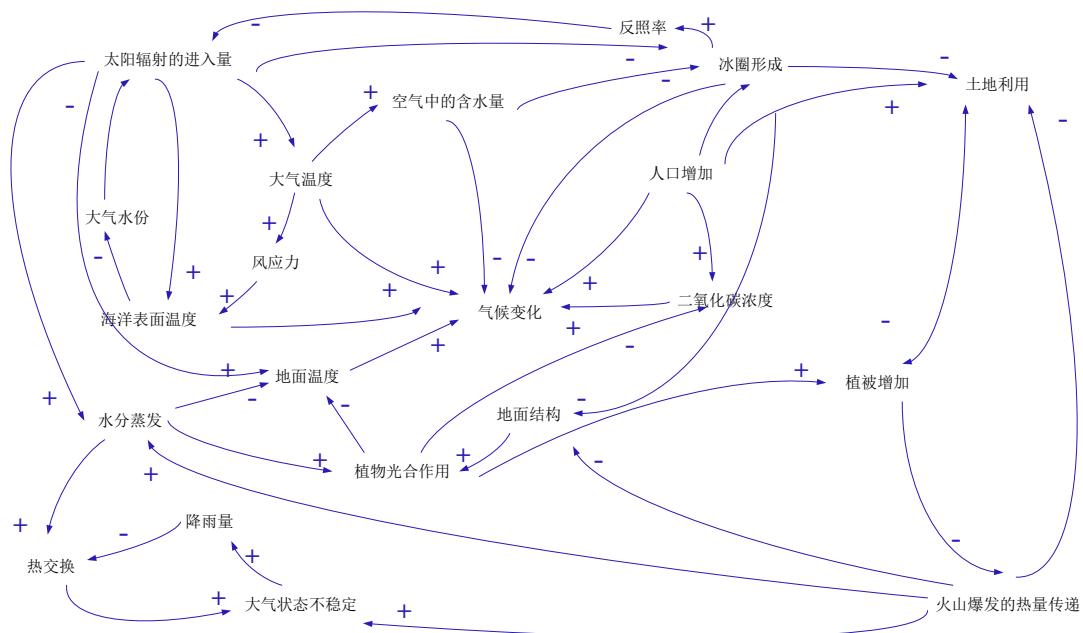


图 6-13 气候变化演化过程总系统流图

6.2.4 如何理解“全球变暖”

全球变暖实质上就是我们生活的大气圈在大气系统内部以及外部的共同作用下，导致吸收和发射的能量相对平衡状态发生略微改变的现象，这种自然状态与人类活动和自然演变规律都有着密不可分的联系。

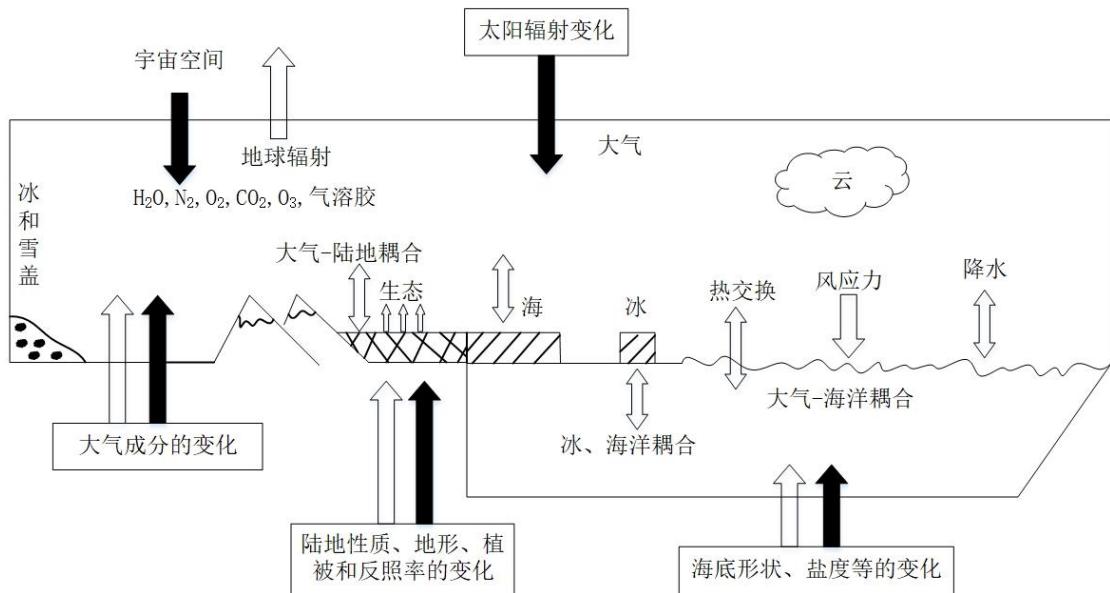


图 6-14 气候变化物质交换示意图

6.3 气候变化对我们的启示

在长时间尺度上，气候变化呈现着一定的规律，这或许是大自然本身存在的规律性事件，但是这是人类发展不能忽略的重要因素。目前来说，气候变化带来的整体结果是“全球变暖”，由此带来的社会环境的变化，比如冰川消融、生态链破坏、恶劣气候频发、热带疾病更大规模的传染等等，是我们需要提前防范的。

“极寒天气”发生的同时，站在另一个角度，我们是否可以考虑“极热天气”的现象？其实，全球大多数人群并不会受到局地极寒（通常是高纬度地区）的影响，生活在中低纬度的他们或许会面临极热气候的影响。甚至有学者认为极热天气的出现将影响人类的死亡率，这或许是我们考虑另一个方向。

“知史以明鉴，查古而知今”。我们需要不断通过了解自然的历史演变过程，发掘新的自然规律。即使越来越多的科学家质疑，碳排放是全球变暖的最主要因素之一，但是站在维持大自然自身能量平衡的角度，人类也不应当过度开发大自然的资源，摒弃“涸泽而渔”的思想。

站在政府组织的角度，应该从源头上遏制生态环境的继续恶化。一方面，要根据温度、水资源、生物等气候环境因子的空间格局与演化趋势，调整生产结构与生活方式；另一方面，积极采取多种减缓措施，坚持把减缓气候变化的核心技术作为优先领域，实施节能优先的能源政策，积极开发可再生能源技术、先进核能技术，风能，太阳能以及高效、洁净、低排放的煤炭利用技术和氢能技术。

站在个人的角度，我们应当响应国家号召，“绿水青山就是金山银山”。积极投身生态文明建设，倡导“低碳出行”，减少温室气体排放，多使用清洁能源，自觉行动维护我们美好的家园。另外，还要关注气象、了解气象、运用气象，提高气象灾害防御的意识。

未来可期！

7. 参考文献

- [1] 孙凤华,杨修群,路爽,杨素英. 东北地区平均、最高、最低气温时空变化特征及对比分析[J], 气象科学,26(2):157-163,2006.
- [2] 张威,纪然.辽宁省地表温度时空变化及影响因素分析[J].生态学报, 18:1-13,2019.
- [3] 韦玮,汪文烈,马林.铜陵地区气温的时空特征分析[J].农业研究与应用,31(04):40-48,2018.
- [4] 谷富,勾晓华,邓洋,苏佳佳,林伟,于爱灵.云贵高原地区 1960-2014 年气温时空变化特征分析.兰州大学学报(自然科学版),54(06):721-730,2018.
- [5] 廖要明 ,陈德亮 ,刘秋锋 . 中国地 气温差时空分布 及 变化趋势 [J]. 气候变化研究 进展,15(04):374-384,2019.
- [6] 丁丽佳. 广东近 50a 来气温时空变化特征[A]. 中国气象学会.第 28 届中国气象学会年会——S4 应对气候变化, 发展低碳经济[C].中国气象学会:中国气象学会,7,2011.
- [7] 陈晶. 黑龙江省气温时空变化特征分析[D].东北农业大学,2013.
- [8] 李柏贞,孔萍,占明锦,章毅之,王怀清.1961—2015 年江西省气温变化特征分析[J].气象与减灾研究,40(03):184-192,2017.
- [9] 海英,高志强.中国百年气候变化及时空特征分析[J].科技通报,26(01):58-62,2010.
- [10] 孟庆佳,林鹏飞,唐晓晖.20 世纪热带海洋海表面温度年际变化的特征[J].海洋科学,39(03):59-65,2015.
- [11] 曾广恩. 东、黄海海表面温度季节内变化的特征研究[D].河海大学,2006.
- [12] 李牧原,罗德海.冬季北极增暖与中纬度环流和极寒天气之间的联系:经向位涡梯度的关键作用[J].中国科学:地球科学,49(09):1335-1345,2019.
- [13] 吴凡. 西北太平洋海表温度年际变化分析[D].中国海洋大学,2013.
- [14] 于文静,余洁,徐凌宇.海表面温时间序列的相关性及复杂性研究 [J]. 计算机技术与发展,29(02):181-184,2019.
- [15] Dang SS, Yang HY, Gao P, et al. A review of research progress on heterogeneous catalysts for methanol synthesis from carbon dioxide hydrogenation[J]. CATALYSIS TODAY, 330:61-75, 2019.
- [16] Ssoussana JF, Lutfalla S, Ehrhardt F, et al. Matching policy and science: Rationale for the '4 per 1000-soils for food security and climate' initiative[J]. SOIL & TILLAGE RESEARCH, 188:3-15, 2019.
- [17] Edwards TL, Brandon MA, Durand G, et al. Revisiting Antarctic ice loss due to marine ice-cliff instability[J]. NATURE, 566(7742):58-, 2019.
- [18] Biskaborn BK, Smith SL, Noetzli J, et al. Permafrost is warming at a global scale[J]. NATURE COMMUNICATIONS, 10, 2019.
- [19] Drappier J, Thibon C, Rabot A, Geny-Denis L. Relationship between wine composition and temperature: Impact on Bordeaux wine typicity in the context of global warming-Review[J]. CRITICAL REVIEWS IN FOOD SCIENCE AND NUTRITION, 59(1):14-30, 2019.
- [20] Frolicher TL, Fischer EM, Gruber N. Marine heatwaves under global warming, 560(7718):360-, 2018.
- [21] Yang K, Yu ZY, Luo Y, et al. Spatial and temporal variations in the relationship between lake water surface temperatures and water quality - A case study of Dianchi Lake[J]. SCIENCE OF THE TOTAL ENVIRONMENT, 624:859-871,2018.
- [22] Dosio A, Mentaschi L, Fischer EM, Extreme heat waves under 1.5 degrees C and 2 degrees C global warming[J]. ENVIRONMENTAL RESEARCH LETTERS, 13(5) ,2018.
- [23] 张任川,张玉臣,刘璟,范钰丹.应用改进卷积神经网络的网络安全态势预测方法[J].计算机工程与应用,55(06):86-93,2019.
- [24] 孙冠华.基于神经网络与非参数核方法 CPI 的 ARMA 预测与非线性改进 [J].统计与决策,34(16):18-21,2018.

- [25] 郑增威,杜俊杰,周燕真,孙霖,霍梅梅.基于多因素稀疏回归预测模型的商家客流量预测[J].计算机应用:1-6,2019.
- [26] 刘彩云,姚俭.基于多因素影响的房地产价格预测模型[J].统计与决策 17:33-38,2017.
- [27] 范可,田宝强,刘颖.2015/2016年极强厄尔尼诺事件下我国动力和统计结合实时气候预测研究[J].大气科学学报,39(06):744-755,2016.
- [28] 赵宗慈,罗勇,黄建斌.10 年时间尺度气候预测[J].气候变化研究进展,11(02):153-156,2015.
- [29] Garnitz J, Lehmann R, Wohlrabe K. Forecasting GDP all over the world using leading indicators based on comprehensive survey data[J]. APPLIED ECONOMICS,51(54):5802-5816,2019.
- [30] De Felice M, Soares MB, Alessandri, A, et al. Scoping the potential usefulness of seasonal climate forecasts for solar power management[J]. RENEWABLE ENERGY, 142:215-223,2019.
- [31] Weisheimer A, Decremer D, Macleod D, et al. How confident are predictability estimates of the winter North Atlantic Oscillation?[J] QUARTERLY JOURNAL OF THE ROYAL METEOROLOGICAL SOCIETY, 145:140-159,2019.
- [32] Centurioni LR, Turton J, Lumpkin R, et al. Global in situ Observations of Essential Climate and Ocean Variables at the Air-Sea Interface[J]. FRONTIERS IN MARINE SCIENCE, 6,2019.
- [33] Ji LQ, Fan K. Climate prediction of dust weather frequency over northern China based on sea-ice cover and vegetation variability[J]. CLIMATE DYNAMICS, 53(1-2):687-705,2019.
- [34] 段青云, 夏军, 缪驰远等. 全球气候模式中气候变化预测预估的不确定性[J]. 自然杂志, 38(3) , 2016.
- [35] 王晓荷, 田茂再. 基于分层贝叶斯时空模型的雾霾天气过程计数分析[J]. 数理统计与管理 06:24-36, 2017.
- [36] 费文绪,伊桑·西格尔.为何全球变暖要为席卷美国的极寒天气负责[J].世界科学 04:26-27,2019.
- [37] 蔡文香. 中国极端气候事件的趋势特征与极值分布[D]. 2016.
- [38] 洪家凤. 阈值模型及其在极端低温中的应用[D]. 扬州大学, 2011.
- [39] 王奕婧. 贝叶斯时空模型在广东省手足口病数据分析中的应用[A]. 中国统计教育学会.2017 年 (第五届)全国大学生统计建模大赛获奖论文选[C].中国统计教育学会:中国统计教育学会,24,2017.
- [40] 邓伟涛,孙照渤.冬季北极涛动与极涡的变化分析[J].南京气象学院学报 05:613-619,2006.
- [41] Lai Zheng,Tarek Sayed. A full Bayes approach for traffic conflict-based before-after safety evaluation using extreme value theory[J]. Accident Analysis and Prevention,131,2019.
- [42] Adam Vaughan. Global warming will drive extreme wet weather in summer[J]. New Scientist,243(3244),019.
- [43] Fei Y, Wang Z, Du YC, et al. Research on rheological and flow behavior of lubricating grease in extremely cold weather[J]. Industrial Lubrication and Tribology,69(6) ,2017.
- [44] 孙建奇,马洁华,陈活泼,汪君,于恩涛,田宝强.降尺度方法在东亚气候预测中的应用[J].大气科学,42(04):806-822,2018.
- [45] 胡娅敏,覃志年,陈丽娟,罗晓玲.基于多时间尺度的回归集成预测模型[J].气象,39(09):1182-1189,2013.
- [46] Y. Xu, V. Ramanathan, D. G. Victor, Global warming will happen faster than we think, Nature, Vol. 564, 30:32, 2018, Dec.
- [47] D. Spratt, I. Dunlop, Existential climate-related security risk: A scenario approach, Policy Paper from Breakthrough-National Centre for Climate Restoration, 2019 May.
- [48] C. Quere, R. Andrew, et al, Global carbon budget 2017, Earth Syst. Sci. Data, 405-448,2018.
- [49] Medhaug, M. B. Stolpe, E. M. Fischer & R. Knutti, Reconciling controversies about the ‘global warming hiatus’, Nature, 2017.
- [50] G. A. Meehl, A. Hu, B. D. Santer & S. Xie, Contribution of the interdecadal Pacific Oscillation to

- twentieth-century global surface temperature tends, *Nature Climate Change*, 6, 1005-1008, 2016.
- [51] X. Chen & K. K. Tung, Global surface warming enhanced by weak Atlantic overturning circulation, *Nature*, 2018.
- [52] T. M. Smith, R. W. Reynolds, A high-resolution global sea surface temperature climatology for the 1961-90 base period, *J. Climate*, 3320-3323, 1998.

8. 附录

EOF 分解的 Matlab 代码

```
clear
close
clc

time=ncread('sst.mnmean.nc','time');
lat=ncread('sst.mnmean.nc','lat');
lon=ncread('sst.mnmean.nc','lon');
sst=ncread('sst.mnmean.nc','sst');
ilat=lat;
nlat=length(ilat);
ilon=lon;
nlon=length(ilon);
itim=1:1985;
ntim0=length(itim);
nyr=ntim0/12;
nph=nlat*nlon;
st1=sst(ilat,ilon,itim);
st2=st1;
ndstp2=length(st2(:,1,1));
ndstp3=length(st2(1,:,:));
nlatlon=ndstp2*ndstp3;
st3=reshape(st2,ntim0,nlatlon);
n1=find(abs(st3(1,:))<999);
npstp=length(n1);
st4=st3(:,n1);
st5=reshape(st4,12,nyr,npstp);
st6_1=st5(1:12,:,:);
st6_2=sum(st6_1,1);
st6=squeeze(st6_2);
clear ntim0;
ntim=nyr;
for i=1:ntim
    st7(i,:)=st6(i,:)-mean(st6,1);
end

st8=reshape(st7,ntim,npstp);
st9=st8;
st10(1:ntim,1:nlatlon)=-999;
st10(1:ntim,n1)=st8;
st11=reshape(st10,ntim,ndstp2,ndstp3);
```

```

for i=1:ndstp2
    st12(:,i,:)=st11(:,i,:)*sqrt(cos(slat(i)*3.14/180));
end

st13=reshape(st12,ntim,ndstp2*ndstp3);
st14(1:ntim,1:npstp)=st13(:,n1);
stp=st14 ;
cst=stp*stp'./ntim;
[eof0,lambda0,explain0]=pcacov(cst);
eof1=stp'*eof0;
for i=1:ntim
    sd=norm(eof1(:,i));
    eof2(:,i)=eof1(:,i)/sd;
end

pca0=stp*eof1;
for i=1:ntim
    sd=norm(pca0(:,i));
    pca(:,i)=pca0(:,i)/sd;
end

lambda=(npstp/ntim)*lambda0;
explain=lambda/sum(lambda)*100;
eof3=eof2';
eof4(1:ntim,1:nlatlon)=-999;
eof4(1:ntim,n1)=eof3;
eof=reshape(eof4,ntim,ndstp2,ndstp3);
for i=1:npstp
    x=[1:ntim]';
    y=stp(:,i);
    nstats = regstats(y,x);%
    stpcoef0(i)=nstats.beta(2);
    stpstat0(i)=nstats.fstat.pval;
end

stpcoef1=stpcoef0;
stpcoef2(1:nlatlon)=-999;
stpcoef2(n1)=stpcoef0;
stpcoef=reshape(stpcoef2,ndstp2,ndstp3);
stpstat1=stpcoef0;
n2=find((stpstat1>0.10));
stpstat1(n2)=-999;
stpstat2(1:nlatlon)=-999;
stpstat2(n1)=stpstat1;

```

```

stpstat=reshape(stpstat2,ndstp2,ndstp3);
fid=fopen(['D:/reg1022.dat'],'w')
B1=stpc coef;
B1=B1';
fwrite(fid,B1,'float32');
B2=stpstat;
B2=B2';
fwrite(fid,B2,'float32');
fclose(fid);

a1=['dset D:/reg1022.dat'];
a2=['undef -999'];
a3=['title Regression trend of MAM NDVI (NDVI/decade:1854-2019'];
a4=['xdef 358 linear 0 2'];
a5=['ydef -88 linear 88 2'];
a6=['zdef 1 linear 0 0'];
a7=['tdef 1 linear 00Z01JAN1854 12mo'];
a8=['vars 2'];
a91=['stpc coef 0 -999 the regression trend of MAM NDVI'];
a92=['stpstat 0 -999 the regression trend of MAM NDVI'];
a10=['endvars'];

fid=fopen('D:/reg1022.ctl','w');

fprintf (fid, '%sn',a1);
fprintf (fid, '%sn',a2);
fprintf (fid, '%sn',a3);
fprintf (fid, '%sn',a4);
fprintf (fid, '%sn',a5);
fprintf (fid, '%sn',a6);
fprintf (fid, '%sn',a7);
fprintf (fid, '%sn',a8);
fprintf (fid, '%sn',a91);
fprintf (fid, '%sn',a92);
fprintf (fid, '%sn',a10);

fclose(fid);
clear a1 a2 a3 a4 a5 a6 a7 a8 a91 a92 a10 B1 B2
fid=fopen(['D:/pca1022.dat'],'w');
for i=1:ntim-
    spca=pca(i,:);
    B=spca;
    fwrite(fid,B,'float32');
end

fclose(fid);

```

```

a1=['dset D:/pca1022.dat'];
a2=['undef -999'];
a3=['title PCA from 185401-201905'];
a4=['xdef 1 linear 1 1'];
a5=['ydef 1 linear 1 1'];
a6=['zdef 1 linear 0 0'];
a7=['tdef 22 linear 00Z01JAN1982 12mo'];
a8=['vars 22'];
a10=['endvars'];

fid=fopen('D:/pca1022.ctl','w');
fprintf (fid, '%sn',a1);
fprintf (fid, '%sn',a2);
fprintf (fid, '%sn',a3);
fprintf (fid, '%sn',a4);
fprintf (fid, '%sn',a5);
fprintf (fid, '%sn',a6);
fprintf (fid, '%sn',a7);
fprintf (fid, '%sn',a8);

for i=1:22
a9=['pca' num2str(i), ' 0 -999 mon mean eof of ndvi'];
fprintf (fid, '%sn',a9);
end

fprintf (fid, '%sn',a10);

fclose(fid);
clear a1 a2 a3 a4 a5 a6 a7 a8 a9 a10 B
fid=fopen(['D:/eof1022.dat'],'w');
for i=1:22
    seof=eof(i,:,:);
    B=squeeze(seof);
    B=B';
    fwrite(fid,B,'float32');
end

fclose(fid);

a2=['undef -999'];
a3=['title EOF from 185401-201905 ','(,num2str(explain(1)),%)'];
a4=['xdef 358 linear 0 2'];
a5=['ydef -88 linear 88 2'];
a6=['zdef 1 linear 0 0'];
a7=['tdef 22 linear 00Z01JAN1982 12mo'];
a8=['vars 1'];

```

```

a9=['eof 0 -999 mon mean eof of ndvi'];
a10=['endvars'];
fid=fopen('D:/eof1022.ctl','w');

fprintf (fid, '%sn',a1);
fprintf (fid, '%sn',a2);
fprintf (fid, '%sn',a3);
fprintf (fid, '%sn',a4);
fprintf (fid, '%sn',a5);
fprintf (fid, '%sn',a6);
fprintf (fid, '%sn',a7);
fprintf (fid, '%sn',a8);
fprintf (fid, '%sn',a9);
fprintf (fid, '%sn',a10);

fclose(fid);
clear a1 a2 a3 a4 a5 a6 a7 a8 a9 a10 B

fid=fopen(['D:/ndvi79_06_mon_mean.dat'],'w');
for t=1:300
    sndvi=ndvi(t,:,:);
    B=squeeze(sndvi);
    B=B';
    fwrite(fid,B,'float32');
end

fclose(fid);

a1=['dset D:/ndvi79_06_mon_mean.dat'];
a2=['undef -999'];
a3=['title NDVI from 185401-201905'];
a4=['xdef 360 linear -179.5 1'];
a5=['ydef 180 linear -89.5 1'];
a6=['zdef 1 linear 0 0'];
a7=['tdef 300 linear 00Z01JAN1854 1mo'];
a8=['vars 1'];
a9=['ndvi 0 -999 mon mean of ndvi'];
a10=['endvars'];
fid=fopen('D:/ndvi1019.ctl','w');

fprintf (fid, '%sn',a1);
fprintf (fid, '%sn',a2);
fprintf (fid, '%sn',a3);
fprintf (fid, '%sn',a4);

```

```
fprintf (fid, '%sn',a5);
fprintf (fid, '%sn',a6);
fprintf (fid, '%sn',a7);
fprintf (fid, '%sn',a8);
fprintf (fid, '%sn',a9);
fprintf (fid, '%sn',a10);

fclose(fid);
clear a1 a2 a3 a4 a5 a6 a7 a8 a9 a10 B
```