



26 **Abstract:** Seasonal change in wind strength and direction is one of the most significant characteristics of  
27 meteorological elements in the monsoon region. In 1999, Lu and Chan defined a monsoon index by using the  
28 seasonal mean value of the South China Sea (SCS) monsoon southerly component intensity, which has a good  
29 correlation with summer precipitation in South China (SC). In order to verify and renovate this index, by using the  
30 monthly mean NCEP/NCAR reanalysis data and the daily precipitation data of the national reference or basic  
31 synoptic stations in SC, we let the wind projected in different directions, not the restricted southerly direction, so  
32 that the performance of the index can be improved. For quantifying the relationship between the SCS's wind monthly  
33 mean projected component intensity in above-mentioned direction and the SC's monthly precipitation, the  
34 Generalized Linear Models and the two forms of the probability distribution, normal distribution and gamma  
35 distribution has been used. The fact of which the SCS's wind projected component intensity in north-west  $11.3^\circ$   
36 have a better correlation with summer precipitation in South China been is discovered. There is a significant  
37 nonlinear relationship between precipitation in Southern China and the new index. The generalized linear model  
38 with Gamma distribution hypothesis for precipitation can obtain better realistic regression results than the normal  
39 distribution hypothesis.

40 **Key words:** South China Sea monsoon; precipitation in South China; nonlinear relationship; generalized linear  
41 models

42

## 43 引言

44 季风作为人类最早总结出的大气现象之一，因为其对温度、降水等天气要素的客观显著  
45 的影响，而一直持续的受到关注。不同时期的学者，借助不同的技术手段，不断地刷新我们  
46 对于季风的认识深度。

47 季风通常被认为是指冬夏之间的风向倒转及气团性质的不同所造成天气现象有季节性  
48 差异的机制，这种不同性质气流的交替往往会带来剧烈的天气变化，如强烈的降水(高辉等，  
49 2001)。南海夏季风由于其地理位置的特殊性，不仅是印度季风和东亚季风连结的媒介，也  
50 是亚澳季风相互联系、相互作用的表现(何金海等，2005)。我国华南地区处于南海季风影  
51 响范围，南海季风异常往往对于华南天气和气候异常有直接影响作用。为了更好的反应南海  
52 夏季风对于我国天气的影响，一系列的季风指数相继提出，例如，施能等(1996)从海陆气  
53 压差角度的研究，孙秀荣和陈隆勋(2000)从海陆热力差异角度的研究，张秀芝等(2002)  
54 从环流特征角度的研究。

55 已有研究表明，华南夏季降水的水汽输送主要有三支，一支是孟加拉湾经青藏高原南侧  
56 通道向东输送到华南地区，另两支分别是中国南海向北输送到华南地区及西太平洋向华南地  
57 区的输送(田红等，2004)。从中国南海向北输送的水汽和西太平洋输送的水汽主要受中国  
58 南海季风强弱的影响，除了直接的水汽输送外，中国南海夏季风的强弱还影响孟加拉湾向东  
59 输送来的水汽的辐合，因此，中国南海夏季风的强弱对华南夏季降水有十分显著的影响作用。

60 Lu 和 Chan (1999) 定义了一种表征中国南海夏季(6-8月)风强弱的指数，其考虑的  
61 主要因子是夏季中国南海地区 1000hPa 层中经向风分量的季节平均强度。事实证明，这个指  
62 数与华南地区逐年的夏季降水总量有很好的相关性。夏季中国南海地区纬向风分量影响西太  
63 平洋支水汽输送且间接地影响孟加拉湾附近自西向东输送来的水汽辐合，从而也对华南地区  
64 夏季降水总量造成一定程度的影响。

65 考虑到上述因素，本文旨在以 Lu 和 Chan 定义的指数为基础，通过综合考虑中国南海  
66 地区夏季的纬向风分量，对已有指数加以改进。这种考虑了纬向风分量的方法，实际是一种  
67 将实际风投影到特定方向上的方法。本文还将对这个风投影分量与华南地区全年逐月降水量  
68 的关系通过建立回归模型的方式进一步定量分析。

69

## 70 1 资料与方法

### 71 1.1 资料

72 本文所采用的资料为中国气象数据共享服务网提供的 1961 至 2011 年中国地面国际交  
73 换站气候资料日值数据集中的降水资料、美国 CISL 研究数据档案室提供的 1961 到 2011 年  
74 NCEP/NCAR 再分析资料的逐月平均风场、高度场资料。需要指出的是,有部分研究认为,  
75 季风活动属季节内尺度范畴,包括季风活跃、中断和季风降水等准双周和周尺度的周期振荡,  
76 本研究中将这些信号当作高斯白噪声,其对应季风强度的方差贡献率较低,但这样的假设是  
77 十分粗糙的。如果要细致的分析不同年份季风的特征,采用月平均资料来定义和描述南海季  
78 风,会滤掉一些有意义的季风活动特征。

79 为保持了与 Lu 和 Chan 1999 年工作的一致性。本文仍选取中南海季风区域为 $[7.5^{\circ} -$   
80  $20^{\circ}\text{N}, 107.5^{\circ} - 120^{\circ}\text{E}]$ , 华南为  $25^{\circ}\text{N}$  以南,  $107.5^{\circ}\text{E}$  以东的陆地区域。华南地区地面气象站  
81 逐日降水资料的记录从 1951 年开始,但华南地区逐日降水资料没有缺测值的站点数上看,  
82 在 1959 年之前,无缺测值的站点较少,而在 1959 年之后,没有缺测值的站点数较多且数量  
83 不再有太大起伏。这大致保证了在本文所选取的研究时间段内,华南地区降水的观测资料受  
84 缺测站点影响而造成的系统误差较小。后文中所涉及的华南月降水总量是指无缺测站降水总  
85 量的算术平均值,且根据联合台风警报中心 (Joint Typhoon Warning Center, JTWC) 的最优  
86 路径资料剔除了受台风影响的日数。

### 87 1.2 方法

#### 88 1.2.1 选取优化风投影方向来综合经向和纬向风分量

89 Lu 定义的表征中国南海夏季风强弱的指数考虑的主要因素是夏季中国南海地区  
90 1000hPa 层中经向风分量  $v$  的季节平均强度,这个指数能很好的反应中国南海夏季风强弱与  
91 华南夏季降水总量的关系。然而为了将中国南海地区夏季的纬向风分量  $u$  对于华南夏季降  
92 水总量的影响加以综合考虑,本文将考虑用风在一个不同于正北方向上的投影,来代替风正  
93 北方向上投影  $v$ 。假设这个方向与正北方向顺时针夹角成  $\alpha$  度,则投影的大小为  $v*\cos \alpha + u*\sin$   
94  $\alpha$ , 即  $(v+u*\tan \alpha)*\cos \alpha$ 。本文考虑的问题是,新指数与华南地区逐年的夏季降水总量的相关  
95 性,将乘数  $\cos \alpha$  略去不会影响两者相关系数的结果。

96 通过在夏季 (6-8 月) 中国南海地区经向风分量  $v$  的基础上加上一定权重  $k$  的纬向风分  
97 量  $u$  来建立改进的指数,以  $v+k*u$  代替  $v$ , 其中  $k$  即为  $\tan \alpha$ , 求取逐年夏季选定区域中  
98  $v+k*u$  的平均值,来讨论其与当年华南地区夏季降水总量的相关性。若有  $k = k^*$  时,  $v+k*u$

99 与华南地区夏季降水总量的相关性最大，则将此时的风投影分量称为优化风投影分量，其符  
100 号记为 $v$ 。

$$101 \quad v \equiv v + k^* * u \quad (1)$$

102 其中 $k^*$ 称为最优权重。

103 为得到最优权重 $k^*$ ，本次研究采取穷举的方案，以 $v+k^*u$ 与华南夏季降水总量相关系  
104 数取得最大值为标准，在合适的解空间中寻找最优的权重解，从而确立适当的权重来对纬向  
105 风分量加以综合考虑，得到改进的指数。解空间的选取主要通过参考 1-12 月逐月平均的纬  
106 向风分量和经向风分量的量级来确定。

107 因为与 Lu 和 Chan 1999 年的文中采取的时间序列长度和起始年份不同，为了更好的反  
108 应改进的指数与原指数在不同时间段与华南地区夏季降水总量相关性的区别，本文从 1970  
109 年开始，取其当年的前 9 年至后 9 年共计 19 年的年份跨度，逐年滑动地求两种指数与华南  
110 地区夏季降水总量的相关系数，从而得到两条相关系数随时间变化的曲线。

### 111 1.2.2 优化风投影分量全年逐月平均值与华南地区降水量的回归分析

112 本文以上内容的基本研究对象是中国南海夏季风强度与华南地区夏季降水总量的线性  
113 相关性。但是，南海季风是华南降水的重要因素，但影响机理复杂，非简单线性关联。在利  
114 用逐月再分析资料定性地给出改进指数所需要的投影方向后，还可以对优化风投影分量与华  
115 南地区降水的非线性关系做进一步定量的分析，这里的非线性变现为一种统计特征，将用到  
116 广义线性回归模型。

117 在实际的研究过程中，在总结优化风投影分量逐月平均值与华南地区逐月降水量之间的  
118 相互关系时，我们发现，从全年逐月的角度看，中国南海季风增强，会使得华南地区降水总  
119 量非线性的增长，这种现象并非只存在于夏季，所以如非特殊说明，本文此后涉及两者定量  
120 关系的“优化风投影分量全年逐月平均值”和“逐月降水量”均指“全年逐月”。另一方面，  
121 以全年逐月南海夏季优化风投影分量为自变量，华南逐月降水强度距平为因变量，将因变量  
122 均匀的划分成几个子区间，各子区间中因变量的标准差随子区间的中值的增大而线性增大，  
123 可以利用最小二乘法回归得到的一个线性关系，这个关系能够通过 95%的信度检验。需要说  
124 明的是，这里采取均匀划分子区间的作法，是出于逐月资料数据量稀疏的考虑，同理于用直  
125 方图近似代替连续变化的函数（Von Mises, 1964）。因此，抛开均匀划分子区间的手段，这  
126 个回归关系就可以理解为逐月降水强度的标准差随南海夏季风强度线性增加，记为  
127  $\sigma = a * v + b$ 。这里引入类似于广义线性回归模型的方法，从而试图更好的刻画中国南海季

128 风强度与华南地区逐月降水量之间的非线性关系。事实上，广义线性回归模型在金融或生物  
129 统计学中都已有广泛运用 (Myers et al. 2012)，而在气候学研究中，也已有十分成功的先例，  
130 如 Gabriele Villarina 等 (2013) 在分析极端降水频率变化时，引入的泊松分布模型下的广义  
131 线性回归方法。广义线性回归模型的连接函数呈指数形式，这也符合降水非负的特性。

132 本文所用到回归模型采用了正态分布和 Gamma 分布两类假设，两类假设条件下回归模  
133 型的具体构建分别如下：

134 正态分布假设：

135 以  $\mu = \exp(\beta_1 * \nu + \beta_0)$  作为连接函数，其中  $\beta_1$  为形状参数， $\beta_0$  为缩放参数， $\mu$  为数学  
136 期望值。并且假设实际值与数学期望值的偏差满足正态分布。依据最大似然方法的概念，样  
137 本总的似然函数表示为：

$$138 \quad P(x_1, x_2, \dots, x_n; \mu_1, \mu_2, \dots, \mu_n; \sigma_1, \sigma_2, \dots, \sigma_n) = \prod_{i=1}^N \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{(x_i - \mu_i)^2}{2\sigma_i^2}\right) \quad (2)$$

139 其中  $N$  表示样本总数； $x_i$  表示第  $i$  个样本元素的值； $\mu_i$  表示第  $i$  个样本元素的数学期望值，  
140 只与样本元素对应的优化风投影分量  $\nu$  月平均值有关，等于  $\exp(\beta_1 * \nu + \beta_0)$ ； $\sigma_i$  表示第  $i$  个  
141 样本元素预期的标准差，只与样本元素对应的优化风投影分量  $\nu$  的月平均值有关，等于  
142  $a * \nu + b$ 。要使其最大，也就是使对数似然函数

$$143 \quad \log P = -\sum_{i=1}^N \frac{(x_i - \mu_i)^2}{2\sigma_i^2} - \sqrt{2\pi} \sum_{i=1}^N \log \sigma_i \quad (3)$$

144 能取到最大值，注意到式 3 右侧第二项对于某一给定的样本为常数，由此问题转化为，  
145 通过改变  $\beta_1$  和  $\beta_0$  的取值，以使得式 4

$$146 \quad f(\beta_1, \beta_0) = -\sum_{i=1}^N \frac{(x_i - \exp(\beta_1 * \nu_i + \beta_0))^2}{2(a * \nu_i + b)^2} \quad (4)$$

147 取到最大值为标准，得到一组确定的最优解  $\beta_1^*$ 、 $\beta_0^*$ 。

148 至此，华南逐月降水总量的数学期望值  $\mu = \exp(\beta_1^* * \nu + \beta_0^*)$ ，和标准差  $\sigma = a * \nu + b$  都  
149 可以表示为随所选南海地区 1000hPa 层中月平均优化风投影分量  $\nu$  变化而变化的函数。随  
150 南海地区 1000hPa 层中月平均优化风投影分量  $\nu$  变化，可以得到一族正态概率分布函数。  
151 这一族正态概率分布函数具有非线性变化的数学期望值。除了关心非线性变化的数学期望  
152 值，我们还关心数学期望值两侧为不同置信度的阈值曲线，以 68.3% 或 95.4% 为例，分别

153 为以数学期望值为对称轴加减一个标准差或两个标准差 (Weisstein, 1999)。这两条阈值线  
 154 所指出的范围, 应该能够包含南海地区 1000hPa 层中月平均优化风投影分量为各个具体值  
 155 时, 华南地区逐月降水总量的较可能出现的情况。

156 Gamma 分布假设:

157 事实上, 降水分布的特征往往不是正态的, 虽然逐月降水总量距平具有的偏度不高, 可  
 158 以用正态分布来近似, 但为了更准确的分析中国南海季风强度与华南地区逐月降水量之间的  
 159 非线性关系, 就需要引入偏态化的模型。有研究表明, Gamma 分布模型是一种适合描述降  
 160 水概率分布模型 (Thom, 1958)。因此本文又以 Gamma 分布假设为出发点构建了回归模型。

161 与正态分布类似, 以  $mode = \exp(\beta_1 * \nu + \beta_0)$  作为链接函数, 其中  $\beta_1$  为形状参数,  $\beta_0$   
 162 为缩放参数,  $mode$  为最大似然值。特别指出的是, 假设最大似然值而不是数学期望值随  $\nu$   
 163 呈指数变化, 是因为这样的假设能够保证 Gamma 分布的参数总是大于 1 的, 从而保证后续  
 164 分析能得到两条连续的阈值曲线。事实上, 在正态分布假设中, 最大似然值总是等于数学期  
 165 望值, 正态分布假设模型的最大似然值也是随  $\nu$  呈指数变化的。依据最大似然方法的概念  
 166 (Choi and Wette, 1969), 似然函数表示为,

167

$$P(x_1, x_2, \dots, x_n; k_1, k_2, \dots, k_n; \theta_1, \theta_2, \dots, \theta_n) = \prod_{i=1}^n \frac{x_i^{k_i-1} \exp(-\frac{x_i}{\theta_i})}{\theta_i^{k_i} \Gamma(k_i)} \quad (5)$$

168 其中  $N$  表示样本总数,  $x_i$  表示第  $i$  个样本元素的值,  $k_i$  表示第  $i$  个样本元素所对应优化风  
 169 投影分量所决定的 Gamma 分布概率密度函数的形状参数,  $\theta_i$  表示第  $i$  个样本元素所对应优  
 170 化风投影分量所决定的 Gamma 分布概率密度函数的比例参数,  $\Gamma(k_i)$  为 Gamma 函数。对数  
 171 似然函数则记为,

172

$$\log P = \sum_{i=1}^N (k_i - 1) \log x_i - \sum_{i=1}^N \frac{x_i}{\theta_i} - \sum_{i=1}^N x_i \log \theta_i - \sum_{i=1}^N \log \Gamma(k_i) \quad (6)$$

173 由于, Gamma 分布的最大似然值  $mode$  等于  $(k-1)*\theta$ , 方差  $\sigma^2$  等于  $k*\theta^2$ 。解二元二  
 174 次方程组得到, 形状参数  $k$  和比例参数  $\theta$  分别为

175

$$\begin{cases} k = 1 + 0.5 * \left[ \left( \frac{mode}{\sigma} \right)^2 + \sqrt{\left( \frac{mode}{\sigma} \right)^4 + 4 * \left( \frac{mode}{\sigma} \right)^2} \right] \\ \theta = \frac{1}{k-1} \left\{ 1 + 0.5 * \left[ \left( \frac{mode}{\sigma} \right)^2 + \sqrt{\left( \frac{mode}{\sigma} \right)^4 + 4 * \left( \frac{mode}{\sigma} \right)^2} \right] \right\} \end{cases} \quad (7)$$

176 这里形状参数  $k$  只取大于 1 的解, 是为了保证 Gamma 分布的最大似然值非零, 从而能  
 177 取到大小两条变化阈值线 (Stacy and Mihram, 1965)。若形状参数取小于 1 的解, 则只能得  
 178 到随南海地区 1000hPa 层中月平均优化风投影分量变化的大值阈值线, 而小值阈值线为非  
 179 变化的零线, 这并不是我们希望得到的模型结构。

180 对于各个样本元素又有,

$$181 \quad \begin{cases} mode_i = \exp(\beta_1 * v_i + \beta_0) \\ \sigma_i = a * v_i + b \end{cases} \quad (8)$$

182  $a$  和  $b$  已由最小二乘法线性回归得到, 将  $\beta_1$  和  $\beta_0$  逐步回带, 直至返回到式 6, 使式 6  
 183 变为  $\beta_1$  和  $\beta_0$  的函数, 表达式冗长, 结果的具体形式此处略。通过改变  $\beta_1$  和  $\beta_0$  的取值, 以  
 184 使式 6 取到最大值为标准, 得到一组确定的最优解  $\beta_1^*$ 、 $\beta_0^*$ 。进一步得到 Gamma 分布假  
 185 设条件下, 随南海地区 1000hPa 层中月平均优化风投影分量  $v$  变化而变化的最大似然值呈对  
 186 数变化的 Gamma 函数, 以及该 Gamma 函数变化的形状参数  $k = k(v; \beta_1^*, \beta_0^*, a, b)$  和缩放  
 187 参数  $\theta = \theta(v; \beta_1^*, \beta_0^*, a, b)$ 。

188 下面需要确定 95%置信度下的阈值线。不同于正态分布假设的情况, 由于 Gamma 函数  
 189 的非对称性, 并不能简单地得到置信度与数学期望值、标准差的关系。但是通过反证法不难  
 190 证明, 两侧阈值需要满足的要求是两阈值对应的概率密度值相等, 这样可以保证这两阈值所  
 191 包括的区间是满足置信度的最窄区间。反证法如下, 假设两阈值所包括区间已经是最窄, 若  
 192 两侧概率密度值不等, 则用概率密度高的一侧的微元代替定积分值相等的概率密度低的一侧  
 193 的微元, 概率密度高的一侧的微元定积分区间宽度一定窄于低的一侧, 运用割补法可以使总  
 194 的定积分区间更窄, 故原假设不成立, 得证。本文没求出这两条阈值线具体的解析形式, 而  
 195 是通过数值计算方法, 求得大小两条阈值线。对于该组变化的  $\theta$  和  $k$ , 高、低阈值  $T_h$ 、 $T_l$  始  
 196 终满足方程组:

$$197 \quad \begin{cases} T_h^{k-1} \exp(-\frac{T_h}{\theta}) = T_l^{k-1} \exp(-\frac{T_l}{\theta}), \quad T_h > T_l \\ \int_{T_l}^{T_h} \frac{x^{k-1} \exp(-\frac{x}{\theta})}{\theta^k \Gamma(k)} dx = 95\% \end{cases} \quad (9)$$

198

## 199 2 结果与讨论

### 200 2.1 选取最优风投影方向来综合经向和纬向风分量的结果与讨论



201 为了确定纬向风分量的权重系数，需要采用穷举方法遍历权重值域。在此之前，为了  
202 确定适合的值得，需要分析纬向风分量和经向风分量之间的关系。

203 图 1 将中国南海地区月平均的纬向风分量和经向风分量分别作为横坐标和纵坐标，以  
204 散点图形式标在图中。由图中所示，要尽量圈取绝大多数散点，宜用长轴和短轴明显不等的  
205 椭圆，而不是用正圆，这说明散点分布呈现明显的偏心特征。也就是所选区域风向风速有明  
206 显季节变化特征，从侧面应征所选区域符合季风区的要求。

207 从图 1 可以看出，在 1000hPa 等压面上，中国南海地区，经向风分量和纬向风分量基本  
208 属于同一个数量级，都是在正负 0 至 10 米每秒的区间范围内，因此，穷举法的权重区间确  
209 定为[-1,1]。

210

211 图 1 南海地区 1-12 月纬向风分量(u)和经向风分量(v)散点图

212 Fig.1 Scatter diagram of monthly west wind components (u) and south (v) over South

213 China sea

214

215 在给定的权重区间中穷举最优权重的结果表现在图 2 上。图中横坐标为穷举区间中权  
216 重系数  $k$ ，纵坐标为求得的不同权重下中国南海地区风投影分量  $v+k*u$  与华南地区夏季降  
217 水总量的相关系数。从图 2 中可以看出，在所选取的  $k$  的值域内，当权重系数  $k$  取-0.2 时，  
218 中国南海地区夏季风投影分量  $v+k*u$  与华南地区夏季降水总量具有比  $k$  取 0 时更好的相  
219 关性，相关系数高达 0.753。也就是说此时  $\tan \alpha$  等于-0.2， $\alpha$  约等于-11.3°。这说明夏季平  
220 均的南海地区 1000hPa 层中的风场，在自正北方向逆时针旋转 11.3° 方向上的投影与华南地  
221 区逐年的夏季降水总量有最好的相关性。所以可以定义出新的指数为，[7.5° - 20°N, 107.5° -  
222 120°E]区域内，1000hPa 层中，南海优化风投影分量  $v \equiv v - 0.2 * u$  的夏季平均值，其中  $v$  为  
223 经向风分量， $u$  为纬向风分量。

224

225 图 2 不同权重 k 下南海地区风投影分量与华南地区夏季降水总量的相关系数 R 图

226 Fig.2 The correlation coefficient (R) between wind projection component over south China  
227 sea and the total summer precipitation in south China under different weight (k)

228

229 图 3 给出的是各年标准化的华南降水异常和指数逐年滑动地两种指数与华南地区夏季  
230 降水总量的相关系数。横坐标是滑动区间的中间值，滑动区间跨度 19 年，因此滑动区间的

231 实际起始年份为横坐标之前的第 9 年。纵坐标是在滑动区间的时间段里两种指数与华南地  
232 区夏季降水总量的相关系数。

233

234 图 3 标准化的华南夏季降水异常和最优风投影分量 (a) 和逐年滑动地两种指数与华南  
235 地区夏季降水总量的相关系数 R 图(虚线: 样本长度为 19 的 95%显著性检验线) (b)

236 Fig.3 Standardized precipitation anomalies s in Southern China and the best wind projection  
237 component (a) and The correlation coefficient (R) between two kinds of indexes and the total  
238 summer precipitation in south China sliding (Dotted line: reference line of 95% significance) (b)

239

240 如图 3a 所示, 标准化的夏季降水与最优风投影分量在 1976 年之前并没有显著的相互关  
241 系, 但在 1976 年之后其正负则较为一致, 尤其在 1987 年之后, 其正负与数值大小都呈现密  
242 切的相互关系。如图 3b 所示, 滑动区间起始年份在 1970 年(即图 3b 中滑动区间中间年份在  
243 1979 年)以前, 两种指数与华南地区夏季降水总量的相关系数并不高, 而在 1970 年以后, 两  
244 种指数分别与华南地区夏季降水总量的相关系数都稳定地维持在较高的水平, 明显高于 95%  
245 置信度的阈值线 (虚线), 且改进的指数始终高于原指数。这说明, 新指数与华南地区夏季  
246 降水总量的相关系数较原指数的改进, 自 1970 年之后始终存在, 与所选的分析时间区间无  
247 关。

248 图 4 和图 5 则分别给出了优化风投影分量在高值年 (1994,1997,2006) 夏季和低值年夏  
249 季 (1980,1989,1996) 1000hPa 等压面上的合成的风场距平和 100hPa 等压面的合成的位势高  
250 度距平场, 此处距平是相对 1979 至 2009 年平均态计算。

251

252 图 4 新指数高值年和低值年夏季 1000hPa 风场合成图 (a 高值年; b 低值年)

253 Fig.4 The composite 1000hPa wind in summer of years during which the new index Present  
254 high value (a) and low value (b)

255

256 图 5 新指数高值年和低值年夏季 100hPa 位势高度合成图 (a 高值年; b 低值年)

257 Fig.5 The composite 100hPa geopotential height in summer of years during which the new index  
258 Present high value (a) and low value (b)

259 从图 4 中可以清楚的看出, 在 $[7.5^{\circ} - 20^{\circ}\text{N}, 107.5^{\circ} - 120^{\circ}\text{E}]$ 范围内, 高值年南风更盛且风  
260 场在  $20^{\circ}\text{N}$  以南的印度洋洋面有显著的西风气流, 这有助于水汽的持续输送; 而在低值年这  
261 个范围内的风场呈现东北向的距平且自印度洋越过中南半岛的水汽输送为负距平。越赤道气  
262 流较强也是优化风投影分量较强年份风场的一个特征。值得注意的是, 新指数高值年往往也  
263 是南亚高压建立偏早的年份 (王黎娟等, 2013)。从图 5 的距平场中可以清晰地看出, 新指  
264 数高值年的夏季南亚高压较低值年偏弱, 南亚高压脊线附近高值年距平场有向北的升度而低  
265 值年有向南的升度, 即南亚高压脊线位置低值年较高值年偏南。已有研究表明, 南亚高压建  
266 立偏早年份里的 4 到 5 月间, 在 $[7.5^{\circ} - 20^{\circ}\text{N}, 107.5^{\circ} - 120^{\circ}\text{E}]$ 的关键区距平风场就已表现出偏  
267 东风特征, 与偏晚年关键区距平风场偏西风相反 (王黎娟和郭帅宏, 2012), 这种偏东风与  
268 偏西风的差异与夏季风场的偏向一致。

269 新指数较旧指数的改进可能存在一些内在机理等待被发觉和研究, 下面将只是探讨性的  
270 提出一种可能的原因。注意到新指数实际上是在  $1000\text{hPa}$  高度上中国南海地区夏季风自正  
271 北方向逆时针旋转  $11.3^{\circ}$  的方向上的投影, 然而在夏季,  $40^{\circ}\text{N}$  向南则盛行南风或西南风  
272 (Yihui Ding, 1994), 因此西偏北方向上的风投影并不直接反应风速率的强弱。但值得注意  
273 的是, 这个方向恰与华南及中南半岛海岸线的走向相垂直, 且偏向青藏高原方向, 在已有的  
274 工作中发现, 南亚高压与南海夏季风强度有显著的关系 (郭帅宏等, 2014)。在这个方向上  
275 的低层风投影量强盛, 也许直接表征出该年海陆间热力差异明显, 在热力差异明显的年份,  
276 水汽将更多的输送到华南地区, 且热力差异对饱和水汽的气团本身的相变提供条件, 从而造  
277 成了华南降水偏多, 反之亦然。

278 表 1 给出了新指数与其他季风指数的对比, 各指数均经过标准化处理, 因此其大小可能  
279 与原始值有所差别 (何金海等, 2001)。

280 从表 1 中可以看出, 新指数在正负情况上与另三种指数基本一致。当另三种指数共同指  
281 出一个强年或弱年时, 新指数也能反应当年的季风强度, 例如 1985、1994 年这样公认的季  
282 风强度强年和 1980 这样的季风强度弱年。而当另三种指数所表示的季风强弱不一致时 (例  
283 如 1984、1988、1992、1993 年), 对应到图 3, 可以看出, 新指数的正负与华南降水异常的  
284 正负具有很好的一致性, 但也有个别例外 (例如 1986 年)。可见, 新指数对于季风的强弱年  
285 判定具有一定参考价值, 而仅从南海季风强度与华南降水异常的定量关系上看, 新指数的参  
286 考价值较其他三种指数更明确。当然, 南海季风的影响十分广泛, 不仅限于华南降水, 所以  
287 在考虑季风强度时, 还是应该综合考虑各个指数的数值, 根据实际情况适当选择。

	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988
施能等	-0.64	-1.93	1.50	0.00	-0.45	0.38	1.08	-1.85	-0.17	1.97
孙秀荣等	0.41	-2.05	0.54	-1.15	-1.00	-0.27	0.44	-0.15	-1.58	0.12
张秀芝等	0.30	-0.73	0.26	0.85	-1.16	-0.16	1.21	0.60	-0.73	-2.09
新指数	-0.39	-1.15	0.10	0.37	-0.6	0.04	1.01	-0.86	-0.11	-0.62

288

(续)

	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
	1.32	1.32	-1.01	0.63	1.19	1.47	1.04				
	-1.44	0.76	-0.82	-0.46	-1.92	0.6	-0.6	-0.52	0.27	-0.83	-1.29
	-0.98	0.87	0.27	-0.55	-0.68	1.41	-1.69	-0.89	0.48	-2.49	0.39
	-1.8	0.16	-0.28	-1.01	1.18	2.03	0.37	-1.26	1.64	-0.35	1.41

289

表 1 新指数与施能、孙秀荣、张秀芝指数的比较

290

Tab.1 The comparison between new index and other indices defined by ShiNeng,

291

SunXiurong, Zhangxiuzhi

292

## 2.2 优化风投影分量全年逐月平均值与华南地区降水量的回归分析的结果与讨论

293

为了更好的反应优化风投影分量与华南降水的定量关系,我们将优化风投影分量分不同  
 294 区间统计了月降水异常的标准差,发现两者有很好的线性相关关系,结果如图 6 所示。在此  
 295 基础上,采用广义线性相关拟合模型对两者的关系进行分析,并且对华南月降水的概率密度  
 296 作了正态分布和 gamma 分布两种近似。

297

298

图 6 优化风投影分量所分各子区间对应华南地区逐月降水量样本标准差及其与子区间  
 299 中值的线性回归线

300

Fig.6 The standard deviation of each sample of monthly rainfall in south China divided by  
 301 split of optimal wind projection component and its regression relationship with mean values of  
 302 wind projection component in each interval

303

304

图 7 为将优化风投影分量逐月平均值均匀划分成不同子区间,其对应的华南地区逐月  
 305 降水量标准差在不同子区间上的直方图,以及优化风投影分量逐月平均值与华南地区逐月降  
 306 水量标准差的线性回归线,线性回归结果通过 95%的信度检验。

307

308 图 7 优化风投影分量全年逐月均值与华南地区全年逐月降水量的回归分析:

309 a 正态分布模型;b Gamma 分布模型

310 Fig.7 The regression analysis of monthly mean of optimal projection weight wind and  
311 monthly precipitation in south China: normal distribution model (a ) and the Gamma distribution  
312 model (b)

313

314 图 7 为优化风投影分量逐月平均值与华南地区逐月降水量的回归分析的结果, 采用的是  
315 是上文提到的最大似然法。a 为正态分布模型的回归结果, 红色实线为华南地区逐月降水量  
316 与优化风投影分量逐月数学期望值的回归曲线, 华南地区逐月降水量随优化风投影分量的增  
317 长呈指数增长; 虚线为降水偏离数学期望值的高低两条阈值线, 阈值线对应的是偏离数学期  
318 望值两个标准差的情况, 置信度为 95%。b 为 Gamma 分布模型的回归结果, 红色实线同图  
319 7a 虚线为降水偏离期望值的高低两条阈值线, 从低阈值(黑色)到期望值(红色)的 gamma  
320 分布概率密度函数积分和从低阈值(黑色)到期望值(红色)的 gamma 分布概率密度函数  
321 积分都等于 0.475。所以华南地区逐月降水量出现再低阈值到高阈值间的置信度也为 95%。  
322 出现在高地阈值线之外的数据点可以视为出现了极端(出现概率低于 5%)的月降水情况

323 由图 7 可以看出, 两种回归模型中, 最大似然值曲线的参数  $\beta_1$  都不是一个接近于零  
324 的值, 这使得在  $v$  的值域[-5.7,3.1]内,  $\beta_1 * v$  不能迅速接近零。也就是说  $\exp(\beta_1 * v + \beta_0)$  并  
325 不能用线性的  $\beta_1 * e^{\beta_0} * v_i + e^{\beta_0}$  来很好的近似, 这反映出华南地区逐月降水量随优化风投影  
326 分量逐月平均值的变化呈现出明显的非线性特征。而从 a, b 图中各自两条阈值曲线的情况  
327 看出, 两条阈值曲线都能包括大多数的散点, 即能很好包括南地区逐月降水量在不同优化风  
328 投影分量逐月平均值的条件下出现的情况。但图 7a 中两条阈值曲线的收束性明显差于 b 中  
329 的两条阈值曲线, 且图 7a 中低值阈值线甚至出现了负值的情况, 这是由于正态分布模型对  
330 于降水资料是一种近似。逐月降水资料本身就具有一定偏度的, 而季节平均降水资料的偏度  
331 则要小许多(张静等, 2013)。而就计算的复杂度讨论, 尽管 Gamma 分布模型能更好的收束  
332 南海地区逐月降水量与不同优化风投影分量逐月平均值的关系, 但两条阈值线需要通过数值  
333 计算, 在“高低阈值对应的概率密度函数相等, 且以其为上下界的概率密度定积分等于置信  
334 度”的条件下求得, 不像正态分布模型, 只需在对应数学期望值基础上加减标准差即可。因

335 此,如果在模型近似程度允许的范围内,建议优先选取正态分布模型;而在近似要求较高的  
336 情况下,建议选取 Gamma 分布模型。

337 事实上,关于降水概率密度符合的偏态分布函数还有很多可行的假设,这里不能穷尽。  
338 即使是对降水概率分布近似度较正态分布假设要高的 Gamma 分布假设也会带来系统性问  
339 题,从图 7b 中可以看出,散落在高阈值线上方的点多于散落在低阈值线下方的点,这可能  
340 是由于 Gamma 分布函数在长尾段和近零段对于降水概率密度的模拟能力不对称而造成的,  
341 这种不足可以通过更换或调整概率分布函数模型加以修正。本文旨在引入模型以揭示华南逐  
342 月降水总量与优化风投影分量的非线性关系。华南逐月降水总量的影响因子还有很多,未来  
343 也会有更多的统计手段被加以运用,希望对于降水这一最基本的天气现象可以有更精确,更  
344 实用的描述。

### 345 3 总结与讨论

346 本研究在 Lu 和 Chan 的 1999 年工作的基础上,去除风在正北方向上作投影的约束条件,  
347 并最终得出结论,夏季中国南海地区 1000hPa 层中实际风在自正北方向逆时针旋转  $11.3^\circ$   
348 的方向上的投影与华南地区夏季降水总量的年际相关性较原指数得到优化。本研究中发现,  
349 在这一方向上的风投影强度与华南逐月降水标准差之间具有很好的线性关系。在此基础上,  
350 利用广义线性回归模型,研究了优化风投影分量的月平均值与华南逐月降水总量的非线性关  
351 系。这一关系不仅包含非线性变化的拟合曲线,还包含其两侧的置信区间阈值线。在概率模  
352 型背景下,本文中旨在广义线性回归和两类降水分布假设背景下,对南海季风强度与华南降  
353 水的关系的研究,比简单分析两者线性相关系数更有价值。从概率模型中不仅能分析出两者  
354 各自的异常,还能监测到两者相互关系异常的现象。被排除在阈值线外的月份,可被清楚的  
355 标记为两者关系异常的月份,从而有助于对南海季风与华南降水之间的关系做更深入的研究。  
356 究。

357 值得注意的是,本文的着眼点是对原有指数 (Lu and Chan, 1999) 的验证和改进,原工  
358 作是对季风年际变化与降水关系这一气候问题进行的论述,因此本文也是立足于年际变化这  
359 一点。但本文也试图从月异常的角度揭示两者的定量关系,研究发现从月异常的角度看,两  
360 者的线性相关性较低,仅为 0.451,这是由于两方面因素造成的,一方面是样本数量的扩大,  
361 另一方面是由于降水偏态分布的特征使得简单的线性并不能很好的反应两者的关系,因此本  
362 文 2.2 小节中采用了广义线性回归来建立两者的定量关系。而对于能反映季风的中断与活跃  
363 的候异常尺度,对两者直接求线性相关,其值仅为 0.213,其原因可能有三点,其中两点与  
364 月异常的情况相同,还有一方面是夏季降水严重受到台风等中尺度系统的影响,尽管可以通

365 过滤波等方式从风场中提取出只与季风有关的指数，但在分析降水时，如果剔除华南受台风  
366 影响的日数，则使得数据在时间上不连续，其对应风场的滤波重构脱离实际物理意义。即从  
367 降水场中剔除台风影响日数和从风场中滤去非季风信号两者不能兼得。综合考虑，本文所得  
368 到的定量结论仅满足于年际和月异常尺度上。

369 本文意在验证和改进 Lu 和 Chan 原有的指数，在验证过程中发现了该指数在上世纪 70  
370 年代前后与华南降水相关性由弱转强的现象。本文如实的呈现了这一现象，但并不能对这一  
371 现象给出明确的解释。关于这一现象的成因，有以下三种猜测：华南降水资料尽管在 1960  
372 年无缺测站点数增多，但数据的均质性在 1978 年前后是否具有一定差异；NCEP/NCAR 再分  
373 析资料在 1979 年是否因为动力框架的更新使得还原中国南海风场的能力增强；全球气候变  
374 化是否导致了南海季风与华南降水相关性的增强。但这都仅仅是猜测，最终的原因可能是其  
375 中之一个或几个、或其他一些尚未被考虑到的因素所综合导致的，对此还需进一步的研究。

376

377 参考文献:

- 378 高辉, 何金海, 徐海明. 2001. 关于确定南海夏季风建立日期的讨论[M]. 南海夏季风建立日期  
379 的确定与季风指数. 气象出版社 : 1-41.
- 380 郭帅宏, 王黎娟, 王苗. 2014. 南亚高压建立早晚与亚洲热带夏季风及中国中东部夏季降水  
381 的关系[J]. 热带气象学报, 30(1): 129-136.
- 382 何金海, 丁一汇, 高辉, 徐海明. 2001. 南海夏季风建立日期的确定与季风指数[M]. 气象出  
383 版社 : 118-123.
- 384 何金海, 徐海明, 周兵, 王黎娟. 2005. 关于南海夏季风建立的大尺度特征及其机制的讨论[J].  
385 气候与环境研究, 5(4): 333-344.
- 386 施能, 朱乾根, 吴彬贵. 1996. 近 40 年东亚夏季风及我国夏季大尺度天气气候异常[J]. 大气  
387 科学, 20(5): 575-583.
- 388 孙秀荣, 陈隆勋. 2000. 东亚海陆热力差指数与中国夏季降水的关系[J]. 南京气象学院学报,  
389 23(3): 378-384.
- 390 田红, 郭品文, 陆维松. 2004. 中国夏季降水的水汽通道特征及其影响因子分析[J]. 热带气象  
391 学报, 20(4): 401-408.
- 392 王黎娟, 郭帅宏. 2012. 4~5 月南亚高压在中南半岛上空建立的年际变化特征及其与亚洲南  
393 部夏季风的关系[J]. 大气科学学报, 35(1): 10-23 .
- 394 王黎娟, 郭帅宏, 何金海, 等. 2013. 4~5 月南亚高压建立早晚年份环流差异及其可能成因  
395 [J]. 大气科学, 37 (6): 1165 - 1178.
- 396 张静, 倪长健, 袁淑杰. 2013. 季度降水分布函数的对比分析——以宜宾市为例[J]. 成都信息  
397 工程学院学报, (1): 29-33.
- 398 张秀芝, 李江龙, 闫俊岳, 等. 2002. 南海夏季风爆发的环流特征及指标研究[J]. 气候与环境  
399 研究, 7(3): 321-331.
- 400 Choi S C, Wette R. 1969. Maximum likelihood estimation of the parameters of the gamma  
401 distribution and their bias[J]. Technometrics, 11(4): 683-690.
- 402 Lu E, Chan J C L. 1999. A unified monsoon index for South China[J]. Journal of Climate, 12(8):  
403 2375-2385.
- 404 Myers R H, Montgomery D C, Vining G G, et al. 2012. Generalized linear models: with applications  
405 in engineering and the sciences[M]. John Wiley & Sons : xi-xiii.
- 406 Stacy E W, Mihram G A. 1965. Parameter estimation for a generalized gamma distribution[J].  
407 Technometrics, 7(3): 349-358.
- 408 Thom H C S. 1958. A note on the gamma distribution[J]. Monthly Weather Review, 86(4): 117-122.
- 409 Villarini G, Smith J A, Vecchi G A. 2013. Changing frequency of heavy rainfall over the central  
410 United States[J]. Journal of Climate, 26(1): 351-357.
- 411 Von Mises R. 1964. Mathematical theory of probability and statistics[M]. New York: Academic  
412 Press :276-289.
- 413 Weisstein E W. 1999. Normal distribution[J]. From MathWorld--A Wolfram Web Resource.  
414 <http://mathworld.wolfram.com/NormalDistribution.html>.
- 415 Yihui D. 1994. The summer monsoon in East Asia[M]. Monsoons over China. Springer Netherlands :  
416 1-90.