

我国绿潮灾害时间序列特征的模拟与预测

刘旭^{1,2,3}, 姜珊¹, 王峥¹, 梁颖祺¹, 何恩业^{1*}

(1. 国家海洋环境预报中心, 北京 100081; 2. 北京林业大学经济管理学院, 北京 100083; 3. 自然资源部海洋灾害预报技术重点实验室 国家海洋环境预报中心, 北京 100081)

摘 要: 基于2010—2019年黄海绿潮卫星遥感影像资料, 构建绿潮覆盖面积时间序列分析方法。将每年5月8日—8月7日成像条件较好的遥感监测数据预处理为周平均时间序列, 将2010—2018年设定为模型训练集, 2019年设定为模型验证集。基于Box-Jenkins法构建了差分整合自回归移动平均模型ARIMA(2, 0, 2)、加法季节性模型SARIMA(1, 0, 0) × (0, 1, 0)₁₂和乘法季节性模型SARIMA(1, 0, 0) × (0, 1, 1)₁₂, 3个模型都通过模型白噪声检验和参数显著性检验, 具有较好的模拟效果和可预测性。SARIMA(1, 0, 0) × (0, 1, 1)₁₂的赤池信息准则值最小, 2019年平均绝对误差为95.56 km², 均方根误差为156.74 km², 与ARIMA(2, 0, 2)相比, MAE提高12%, RMSE下降1.2%, SARIMA(1, 0, 0) × (0, 1, 0)₁₂的预测精度最低, MAE和RMSE分别为115.12 km²和192.16 km²。

关键词: 绿潮; 差分整合自回归移动平均模型; 时间序列法; 遥感监测

中图分类号: X55 **文献标识码:** A **文章编号:** 1003-0239(2023)02-0056-11

0 引言

绿潮通常是指由于大型绿藻(如浒苔)的过度增殖、生长和聚集而形成的一种生态灾害^[1]。绿潮主要分布于南北半球的温带海域, 其中亚得里亚海、波罗的海、法国布列塔尼附近海域及我国黄海海域的绿潮问题比较突出^[2-3], 例如2008年我国青岛近岸海域爆发了规模极大的绿潮灾害^[4-5]。绿潮灾害不仅影响景观环境、海运交通和水上活动^[6-7], 还会对海草床生境、底栖生物群落和水产养殖造成一定威胁^[8]。在全球气候变化和近岸海域富营养化的背景下, 绿潮灾害受到了政府和学术界的高度关注^[9]。自2007年我国首次爆发大规模绿潮以来, 山东省和江苏省沿海地方政府每年夏季都投入大量人力、物力对浒苔进行打捞收集和处理^[2]。2016年, 青岛海洋科学与技术试点国家实验室设置了“近海生态灾害发生机理与防控策略”鳌山科技计划项

目, 重点针对黄海绿潮问题进行联合攻关。2021年, 我国黄海浒苔覆盖面积出现历史最大值, 为应对绿潮灾害, 青岛市紧急启动“近岸海域打捞浒苔应急处置”工作, 江苏省也立即拨付了浒苔绿潮联防联控省级工作经费。

绿潮预测预警研究有助于在绿潮来临之前为政府部门提供有效的决策辅助信息, 降低绿潮灾害的影响程度和损失, 并为提升我国绿潮防范减灾的国际影响力提供技术支撑^[10-11]。绿潮生消变化受水文气象、海水富营养化等多重因素影响, 爆发机制及各因素之间的耦合作用尚不清楚^[5], 因此绿潮预测是绿潮研究领域的热点和难点^[2]。从物理海洋学研究方向来看, 绿潮预测可以结合数学模型建立生态动力学模式进行。自然资源部第一海洋研究所利用中国近海三维海浪-潮流-环流耦合数值模式对黄海浒苔的漂移开展模拟和预测^[12]。自然资源部北海预报中心基于三维全动力POM(Princeton Ocean

收稿日期: 2021-08-25; 修回日期: 2022-03-01。

基金项目: 国家自然科学基金(41976221, 41576029); 国家重点研发计划“地球观测与导航”重点专项资助项目(2021YFB3900405)。

作者简介: 刘旭(1986-), 女, 工程师, 博士, 主要从事海洋灾害评估及系统分析工作。E-mail: fairyjujube@126.com

*通信作者: 何恩业(1981-), 男, 高级工程师, 学士, 主要从事海洋生态环境、生态灾害和水文气象预报研究。E-mail: heenye@163.com

Model)海洋模式,利用多源遥感影像监测数据,构建了黄海绿潮监测预警综合业务化系统^[13]。国家海洋环境预报中心于2008年构建了浒苔漂移路径预测系统^[14],2019年系统耦合了浒苔生长消亡过程的生态模块^[15],并基于中分辨率成像光谱仪(Moderate-resolution Imaging Spectroradiometer, MODIS)遥感卫星对当年浒苔的生消过程进行了分析^[16]。从数据分析研究方向来看,长期的绿潮监测数据积累为大数据分析提供了数据支撑,目前主要采用数据挖掘和机器学习技术对绿潮规模和影响因子进行分析。唐莹莉^[17]基于支持向量机法分析了海温、天气现象、风向、风力和浪高因素对不同时期绿潮分布面积的影响程度并建立了预测模型,认为海温是绿潮监测中的首要因子。何世钧等^[18]根据主导因子的权重交点分析了海温、浪高、雾霾对绿潮扩散影响的变化规律,并对2015年进行了预测,误差仅为1.8%~7%。刘赞玥^[11]采用灰色粗糙算法优化的误差反向传播神经网络(Back-Propagation Neural Network, BP)方法对绿潮覆盖面积进行了长期预测,解决了传统BP神经网络算法只限于小规模数据范围的预测问题,该方法为绿潮防治提供了辅助性信息。

目前,基于数据驱动的分析方法预测绿潮仍处于起步和探索阶段。利用时间序列分析法对历史数据的时间特征和变化进行分析,可有效提取结构性成分和不规则变动成分,为定量分析和预测综合因素影响下绿潮的生消变化提供了可能^[19]。近年来,时间序列分析法已用于赤潮的长期预测^[19-22],但用于研究绿潮生消变化的研究鲜有报道。本文基于2010—2018年我国黄海多源卫星遥感反演的绿潮覆盖面积数据,利用Box-Jenkins建模方法,构建了差分整合自回归移动平均(Autoregressive Integrated Moving Average, ARIMA)模型、加法季节性差分整合自回归移动平均(Seasonal Autoregressive Integrated Moving Average, SARIMA)模型和乘法SARIMA模型,通过与2019年实测值进行比较并计算了模型的预测精度,旨在为绿潮灾害的长期预测提供技术支撑。

1 数据与方法

1.1 数据来源与预处理

黄海绿潮覆盖面积原始数据来源于2010—

2019年《国家卫星海洋应用中心绿潮遥感监测通报》中每日业务化绿潮遥感影像反演结果,区域范围为33°~38°N, 119°~123°E,其中“海洋一号B”卫星(HY-1B)和MODIS卫星的分辨率为250 m,“高分三号”卫星(GF-3)、“高分四号”卫星(GF-4)和“海洋一号C”卫星(HY-1C)的分辨率为50 m^[15]。虽然每日业务化开展绿潮遥感反演监测,但由于云层对光学卫星遥感成像条件会有限制,难以获得每日有效数据。因此,按照绿潮生消的主要起始时间,将每年5月8日—8月7日设定为年度绿潮分析周期,以星期为单位将每月分成1—7日、8—14日、15—21日、22日—月底4个时段,挑选每7 d周期内云量较少或无云情况下获取的绿潮遥感影像数据,将有效数据进行平均并作为绿潮覆盖面积的各星期合成值,绿潮生消年度周期设定为12周,将数据预处理为等时间间隔的合成序列。2010—2018年数据作为训练集用于构建模型,2019年数据作为测试集用于模型预测检验,评估模型预测精度。

1.2 研究方法

时间序列分析法是根据一组相依有序的离散数据,建立反映时间序列中所包含的动态依存关系的数学模型,并进行未来状态预测的数据驱动型分析方法^[23]。利用ARIMA模型构建非季节性绿潮覆盖面积预测模型,将差分过程、自回归模型和平均移动模型相结合,ARIMA(p, d, q)模型可表示为:

$$y_t = \Delta^d x_t \quad (1)$$

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \cdots + \alpha_p y_{t-p} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \cdots + \beta_q \varepsilon_{t-q} \quad (2)$$

式中: x_t 为当前时刻的原始值; y_t 为 d 次差分后的测量值; Δ^d 为 d 阶差分; $\varepsilon_t, \varepsilon_{t-1}, \cdots, \varepsilon_{t-q}$ 为当前时刻和前期若干时刻的随机误差项; $\alpha_1, \alpha_2, \cdots, \alpha_p$ 为待确定的自回归系数; $\beta_1, \beta_2, \cdots, \beta_q$ 为待确定的平均移动系数^[23]。

季节性时间序列是指序列经过固定时间间隔内呈现出相似的波动规律,时间序列可被分解为趋势(T)、季节性(S)和随机误差(I)^[23],可分别采用加法季节性模型SARIMA(p, d, q) \times (0, 1, 0)_s与乘法季节性模型SARIMA(p, d, q) \times (P, D, Q)_s来对比模型的模拟精度与预测精度,分析季节性影响对ARIMA成分的影响,并筛选出绿潮覆盖面积最优

预测模型。

加法模式:

$$Y = T + S + I \quad (3)$$

模型构造如下:

$$\Delta_s \Delta^d x_t = \frac{L(B)}{A(B)} \varepsilon_t \quad (4)$$

$$\begin{cases} A(B) = 1 - \alpha_1 B - \alpha_2 B^2 - \cdots - \alpha_p B^p \\ L(B) = 1 - \beta_1 B - \beta_2 B^2 - \cdots - \beta_q B^q \end{cases} \quad (5)$$

乘法模式:

$$Y = T \times S \times I \quad (6)$$

模型构造如下:

$$\Delta_s^D \Delta^d x_t = \frac{L(B)V(B^s)}{A(B)U(B^s)} \varepsilon_t \quad (7)$$

$$\begin{cases} U(B^s) = 1 - \tau_1 B^s - \tau_2 B^{2s} - \cdots - \tau_p B^{ps} \\ V(B^s) = 1 - \theta_1 B^s - \theta_2 B^{2s} - \cdots - \theta_q B^{qs} \\ \Delta^d = (1 - B)^d \\ \Delta_s^D = (1 - B^s)^D \end{cases} \quad (8)$$

式(3)一(8)中: S 为周期步长; B 为滞后算子; p 、 P 、 q 、

Q 分别表示非季节项和季节项的自回归与移动平均特征的最大滞后阶数; Δ^d 、 Δ_s^D 分别表示 d 阶非季节性差分 and D 阶周期步长为 S 期的季节差分; ε_t 为随机误差项; $A(B)$ 、 $U(B^s)$ 分别表示非季节性和季节性自回归多项式; $L(B)$ 、 $V(B^s)$ 分别表示非季节性和季节性移动平均多项式;其余参数含义同式(2)。

采用R语言,基于Box-Jenkins建模方法建立时间序列模型。关键步骤见图1。

2 结果与讨论

2.1 时间序列构建

2010—2019年,国家卫星海洋应用中心绿潮遥感监测系统每年首次监测到绿潮的时间主要集中于5月中下旬,其中2018年最晚(5月31日),2011年也偏晚(5月27日),2012年、2015年和2016年均于5月16日,因此5-2nd(5月第2周,即5月8—14日,以此类推)或5-3rd数据存在0值(见图2)。绿潮消亡

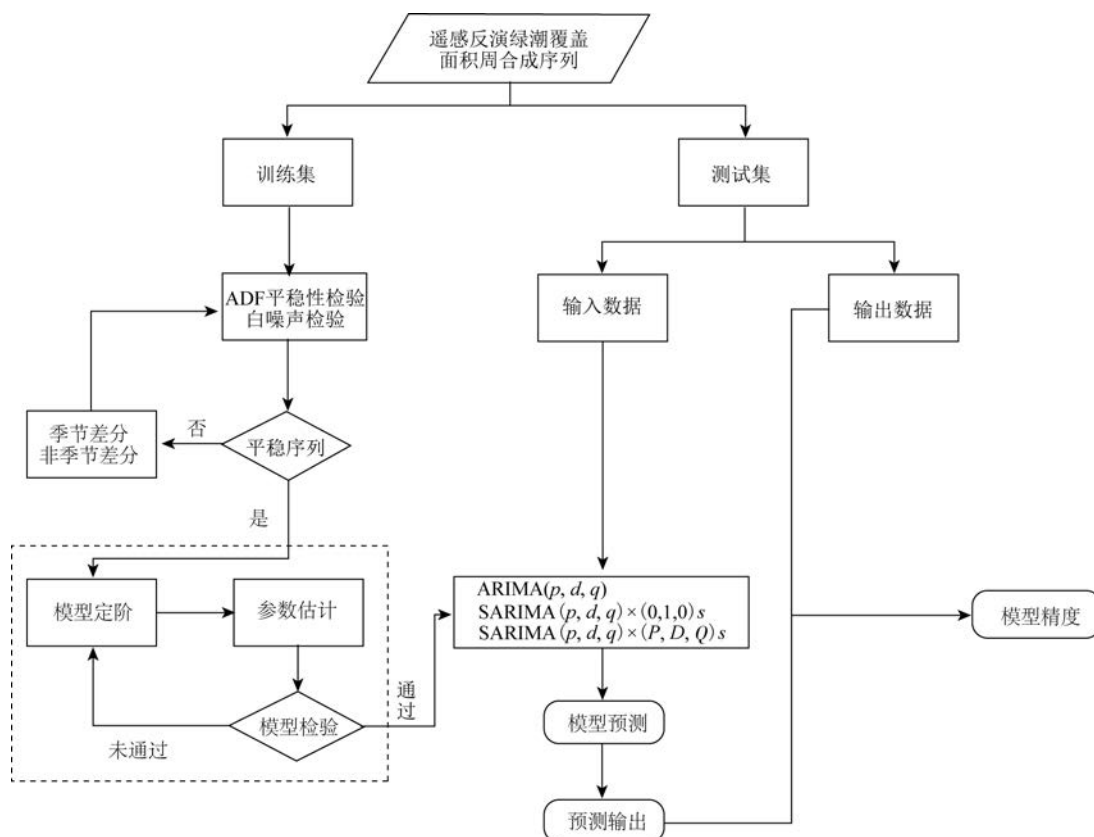


图1 绿潮覆盖面积时间序列预测建模流程图

Fig.1 The flowchart of time series prediction modeling of green tide coverage area

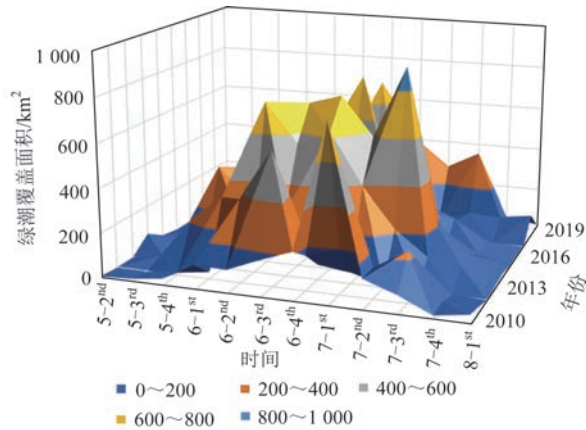


图2 2010—2019年卫星遥感监测的黄海绿潮周平均覆盖面积分布

Fig.2 Weekly averaged coverage area of green tide monitored by satellite remote sensing from 2010 to 2019 in the Yellow Sea

结束时间集中于7月下旬—8月中旬,其中2017年结束最早(7月13日),其次为2012年和2016年,分别为7月29日和7月28日。因此,2017年7月后两周和8月第1周的数据为0,2012年和2016年8月第1周数据为0。2010—2019年每年黄海绿潮灾害遥感监测持续时间为65~121 d,平均持续时间为84 d,其

中2017年最短,为65 d,2019年最长,为121 d。

2012年绿潮到达峰值时间最早,为6-1st,绿潮面积为330 km²,为10 a中绿潮规模峰值的最小值。2014年绿潮到达峰值时间最晚,为7-2nd,绿潮面积为905 km²,为10 a中绿潮规模峰值的最大值。其余年份中,绿潮规模达到峰值时间段主要集中于6-3rd和6-4th,峰值面积平均为576 km²。

2.2 序列平稳性检验

将2010—2018年周平均绿潮覆盖面积作为训练集构建模型。从时间序列图中可以看出原始序列呈明显的年度季节波动特点(见图3a),自相关系数(Auto-Correlation Function, ACF)在滞后期(Lag)为12时出现明显的正峰值(见图3b),也说明序列存在季节波动规律。将原始序列进行ADF(Augmented Dickey Fuller)平稳性检验和白噪声检验,结果表明序列为平稳非白噪声序列(见表1),ADF单位根检验法季节波动并未影响序列的平稳性。因此,根据Box-Jenkins建模要求可直接构建原始序列的自回归滑动平均模型(Autoregressive Moving Average Model, ARMA)。

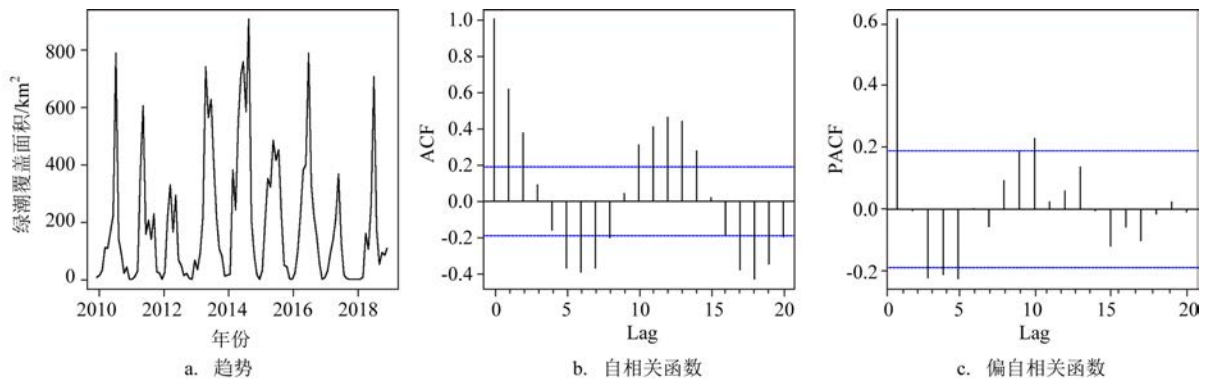


图3 原始序列平稳性检验图

Fig.3 Original series stationarity test

表1 ADF检验和Q检验统计结果

Tab.1 Results of ADF test and Q residual test

类别	ADF 检验		Q 检验		结论
	统计量	p 值	统计量	p 值	
原始序列	-6.071 5	0.01	172.82	2.2×10^{-16}	平稳非白噪声
季节差分	-3.313 9	0.07	40.33	6.33×10^{-5}	非平稳非白噪声
1阶季节差分	-6.667 5	0.01	44.226	1.39×10^{-5}	平稳非白噪声

注:R软件ADF检验中,当p值小于0.01时统一显示为0.01。

采用季节差分法($S=12$)可以消除年度循环季节效应。图4a表明通过步长为12的季节差分基本消除了季节效应,ACF呈现出指数回落特征(见图4b),偏自相关系数(Partial Auto - Correlation Function, PACF)呈现出截尾特征并在第12阶出现峰值(见图4c),说明通过差分可以平滑序列的季节波动特征,但由于仍存在短期效应,所以仍为非平稳序列(见表1)。

为构造平稳序列再次对数据进行1阶差分,获得的时间序列图更加平稳(见图5a),但与图4a相比,图5a也损失了序列的部分细节信息。ACF迅速截尾(见图5b),PACF呈现出指数下降规律(见图5c),并均在第12阶出现明显峰值,经过ADF检验和Q检验,该序列为平稳非白噪声序列(见表1)。

2.3 模型定阶

2.3.1 ARMA模型

原始序列经过ADF检验后为平稳非白噪声序

列。将2010—2018年的绿潮覆盖面积周平均值序列直接进行ARIMA训练,差分阶数 $d=0$ 。根据序列的ACF和PACF初步确定模型的阶数,由于ACF呈现出明显的拖尾状态(见图3b),PACF呈现出不太明显的截尾状态(见图3c),因此尝试采用自回归(AutoRegressive, AR)模型或ARMA模型进行模型初步定阶。依据赤池信息准则(Akaike Information Criterion, AIC)进行最优拟合模型评价,当 $p=2, q=2$ 时, AIC取得最小值为1 409.76,故最优非季节性模型初步设定为ARIMA(2,0,2)。

2.3.2 加法季节性模型

对原始序列进行1阶12步差分构建平稳序列(见表1),提取出季节信息和趋势信息后,用ARIMA模型进行拟合。差分后的ACF呈现1阶明显的截尾特征,除了1阶、11阶、12阶、13阶大于2倍标准差外(见图5b),其他阶数的自相关系数基本都在2倍标准差范围内波动。PACF呈现出下滑轨迹,1阶、2阶、11阶、12阶明显大于2倍标准差(见图

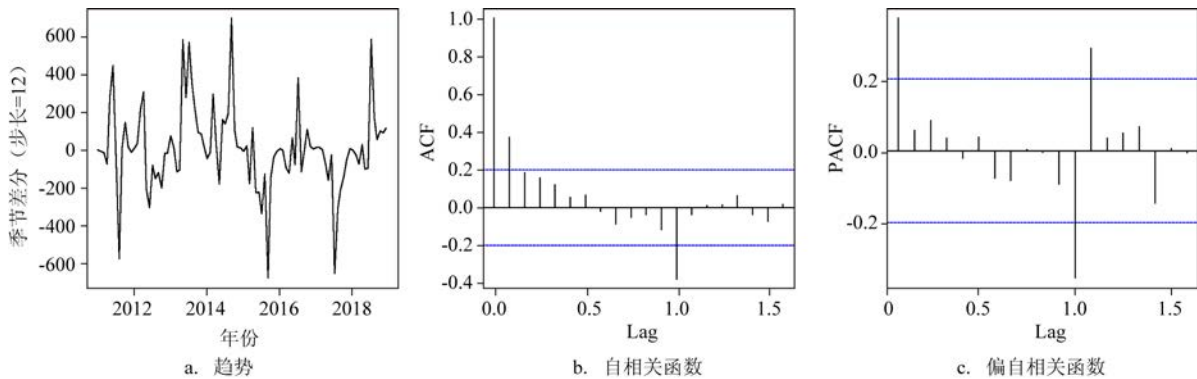


图4 季节差分平稳性检验图

Fig.4 Seasonal differential series stationarity test

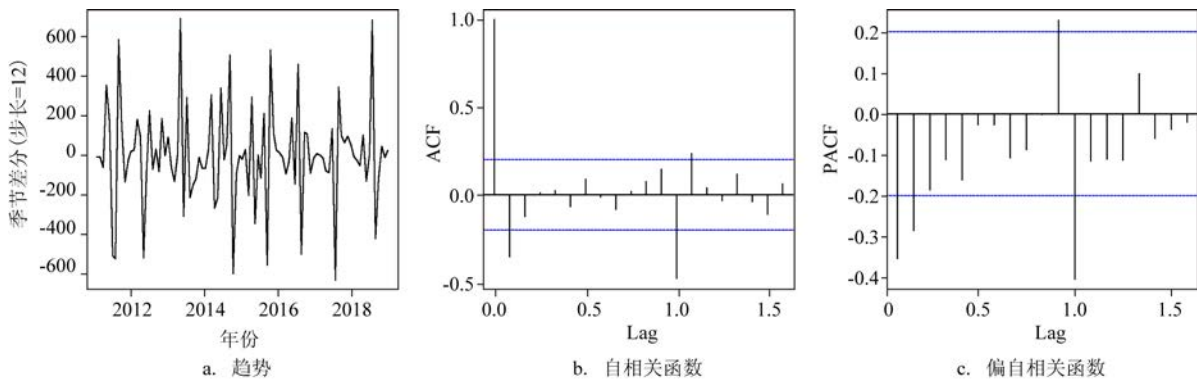


图5 1阶季节差分平稳性检验图

Fig.5 First-order seasonal difference series stationarity test

5c),因此尝试采用滑动平均(Moving Average,MA)模型或ARMA模型进行模型初步定阶。经计算(见表2),当 $p=1, q=1$ 时,AIC取得最小值1 298.97,故最优加法SARIMA模型初步设定为 $SARIMA(1,0,1) \times (0,1,0)_{12}$ 。

2.3.3 乘法季节性模型

根据ACF和PACF的特点,且图5b呈现出在1阶迅速截尾、在12阶出现峰值的特点,因此设定 $P=0$;图5b呈现出1阶大于2倍标准差、2阶接近2倍标准差、3阶后截尾、12阶出现峰值的特点,因此设定 $Q=1$ 或 $Q=2$ 。经计算(见表3),当 $p=1, d=0, q=1, P=1, D=1, Q=1$ 时,AIC取得最小值1 270.32,故最优乘法SARIMA模型初步设定为 $SARIMA(1,0,1) \times (1,1,1)_{12}$ 。

2.4 模型参数的估计及检验

在采用极大似然法对初步定阶的模型进行参数估计的基础上,为避免出现过拟合问题,我们对参数进行t检验。在以上3个初步定阶模型中只有 $ARIMA(2,0,2)$ 通过参数显著性t检验(见表4)。 $SARIMA(1,0,1) \times (0,1,0)_{12}$ 中MA1系数未通过参数显著性检验,因此重新采用加法SARIMA模型进行训练;令 $p=1$,参数通过显著性检验,AIC=1 297.5;

表2 加法SARIMA模型模拟精度

Tab.2 Simulation accuracy of additive SARIMA model

非季节项参数		季节项参数 $(0,1,0)_{12}$
$p=0$	$q=1$	1 299.71
	$q=2$	1 300.36
$p=1$	$q=1$	1 298.97
	$q=2$	1 300.61
$p=2$	$q=1$	1 300.73
	$q=2$	1 302.55

表3 乘法SARIMA模型模拟精度

Tab.3 Simulation accuracy of multiplication SARIMA model

非季节项参数		季节项参数			
		$(1,0,1)$	$(1,0,2)$	$(1,1,1)$	$(1,1,2)$
$p=0$	$q=1$	1 416.79	1 416.92	1 276.00	1 277.56
	$q=2$	1 413.49	1 414.52	1 274.17	1 276.1
$p=1$	$q=1$	1 406.58	1 406.61	1 270.32	1 274.16
	$q=2$	1 407.5	1 407.67	1 272.32	1 274.16
$p=2$	$q=1$	1 407.33	1 407.43	1 272.32	1 274.08
	$q=2$	1 409.09	1 408.26	1 272.99	1 276.16

表4 参数显著性估计与模型残差检验结果

Tab.4 Results of parameter significance estimation and model residual test

模型	模型参数	参数标准误差	参数显著性检验(p值)	模型Q检验(p值)	结论
ARIMA(2,0,2)	AR1=1.69	AR1=0.03	AR1:0.00	0.36	通过
	AR2=-0.95	AR2=0.03	AR2:0.00		
	MA1=-1.43	MA1=0.13	MA1:0.00		
	MA2=0.77	MA2=0.12	MA2:0.00		
	C=190.77	C=19.09	C:0.00		
加法SARIMA $(1,0,1) \times (0,1,0)_{12}$	AR1=0.57	AR1=0.25	AR1:0.01		未通过参数检验
	MA1=-0.25	MA1=0.31	MA1: 0.21		
加法SARIMA $(1,0,0) \times (0,1,0)_{12}$	AR1=0.36	ARI=0.09	AR1:0.00	0.36	通过
乘法SARIMA $(1,0,1) \times (1,1,1)_{12}$	AR1=0.72	AR1=0.14	AR1:0.00	0.36	未通过参数检验
	MA1=-0.36	MA1=0.20	MA1:0.03		
	SAR1=-0.09	SAR1=0.19	SAR1:0.31		
	SMA1=-0.67	SMA1=0.22	SMA1:0.00		
乘法SARIMA $(1,0,0) \times (0,1,1)_{12}$	AR1=0.42	AR1=0.09	AR1:0.00	0.36	通过
	SMA1=-0.71	SMA1=0.13	SMA1:0.00		

$p=2$ 时,AR2系数不显著,未通过参数显著性检验。在乘法SARIMA(1,0,1)×(1,1,1)₁₂中,SAR1系数未通过参数显著性检验,根据检验结果令 $q=0, P=0$,重新对模型重新定阶(见表5),其中只有SARIMA(1,0,0)×(0,1,1)₁₂通过参数显著性检验,AIC=1 269.02(见表4)。

为避免出现欠拟合问题,对通过参数显著性检验的3个模型分别进行残差白噪声Q检验(见图6)。结果表明,ARIMA(2,0,2)模型、SARIMA(1,0,0)×(0,1,0)₁₂模型、SARIMA(1,0,0)×(0,1,1)₁₂模型均通过模型残差检验(见表4),说明模型均有效提取了序列中的信息,均具有可预测性。

表5 乘法SARIMA模型重新定阶AIC值

Tab.5 Reorder AIC values of multiplicative SARIMA model

非季节项参数		季节项参数			
		(0,0,1)	(0,0,2)	(0,1,1)	(0,1,2)
$q=0$	$p=1$	1 428.47	1 421.47	1 269.02	1 270.93
	$p=2$	1 430.33	1 422.99	1 269.04	1 270.77

虽然3个模型均通过了Ljung-Box残差检验,但是模型残差自相关系数结果表明,当ACF滞后期为0.5时,ARIMA(2,0,2)的模型残差略大于两倍标准差,表明还有短期效应信息未提取出来(见图6a);

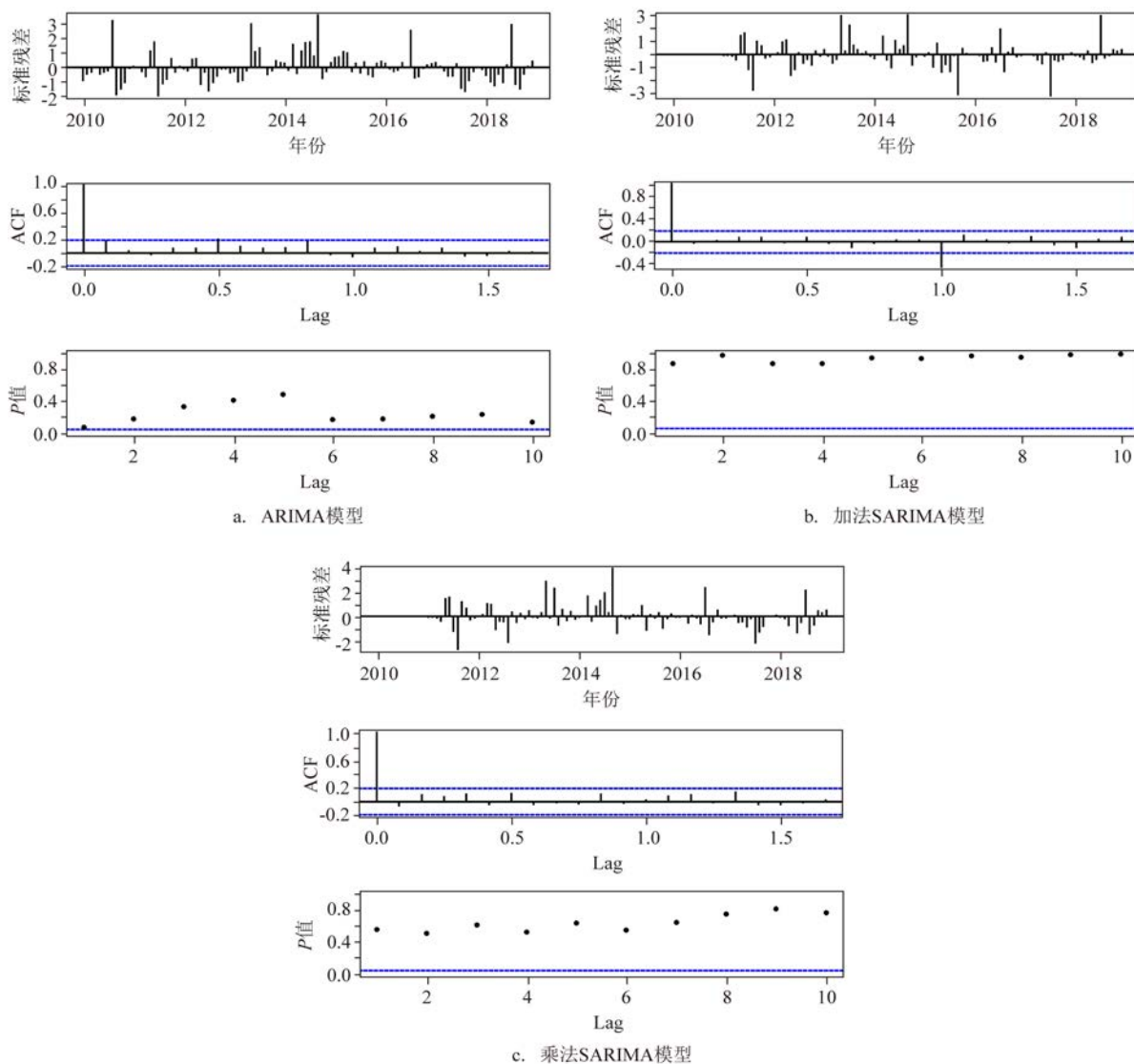


图6 模型残差检验结果

Fig.6 Model residual test results

当ACF滞后期为1.0时,加法SARIMA(1,0,0)×(0,1,0)₁₂的模型残差显著大于两倍标准差,表明该模型将差分固定分配给季节项,但未完全将长期季节波动信息提取出来(见图6b);SARIMA(1,0,0)×(0,1,1)₁₂的模型残差均小于两倍标准差,说明其短期效应和季节效应均得到有效提取(见图6c)。根据AIC评估模型的模拟精确度(见表2和表5),结果表明乘法季节性模型SARIMA(1,0,0)×(0,1,1)₁₂的AIC值最小,说明其模拟效果优于其他两个模型。

2.5 模型预测结果与精度对比分析

采用趋势外推法,依据以上3个通过参数显著性检验和残差检验的模型,分别对2019年绿潮覆盖面积进行超前12步预测(见图7)。需要说明的是,绿潮覆盖面积不可能是负值,图中置信区间仅为统计计算结果。在原始序列构建的模型中,2019年绿潮最大覆盖面积低于实测值(见图7a),经过1阶季节差分构建的两个模型峰值的预测值高于实测值(见图7b和7c),其中加法SARIMA模型的峰值预测结果与实测值偏差较大。ARIMA(2,0,2)模型峰值预测时间为6-3rd(6月15—21日)、SARIMA(1,0,0)×(0,1,0)₁₂模型和SARIMA(1,0,0)×(0,1,1)₁₂模型峰值预测时间均为6-4th(6月22—30日),2019年绿潮到达峰值的时间为10 a中偏晚,为7-1st(7月1—7日),因此加法SARIMA模型和乘法SARIMA模型与实测结果较为接近。

采用平均绝对误差(Mean Absolute Error, MAE)和均方根误差(Root Mean Squared Error, RMSE)评价模型的预测精度(见表6)。SARIMA(1,0,0)×(0,1,1)₁₂的MAE最小,ARIMA(2,0,2)的RMSE最小,SARIMA(1,0,0)×(0,1,0)₁₂的MAE和RMSE均高于其他两个模型。SARIMA(1,0,0)×(0,1,1)₁₂的MAE低于ARIMA(2,0,2),但RMSE高于ARIMA(2,0,2),这可能由于RMSE进行了平方计算,因此放大了异常值(峰值)对预测精度的影响^[23]。ARIMA(2,0,2)预测出的2019年绿潮最大覆盖面积为309 km²,SARIMA(1,0,0)×(0,1,1)₁₂的预测结果为502 km²,遥感实测值为389 km²。仅对峰值1期进行误差分析,ARIMA(2,0,2)和SARIMA(1,0,0)×(0,1,1)₁₂的绝对误差分别为90 km²和113 km²,由于RMSE的计算方式放大了峰值对整个2019年12周预测精确度的影响,可能导致了MAE和RMSE预测评价指向了不同模型。综合评价两个模型的预测精度,SARIMA(1,0,0)×(0,1,1)₁₂的MAE较ARIMA(2,0,2)提高了12%,RMSE仅下降了1.2%,预测的峰值出现时间与实测值更为接近。

表6 模型预测精度评价

Tab.6 Model prediction accuracy evaluation

预测模型	MAE/km ²	RMSE/km ²
ARIMA(2,0,2)	108.74	154.87
SARIMA(1,0,0)×(0,1,0) ₁₂	115.12	192.16
SARIMA(1,0,0)×(0,1,1) ₁₂	95.56	156.74

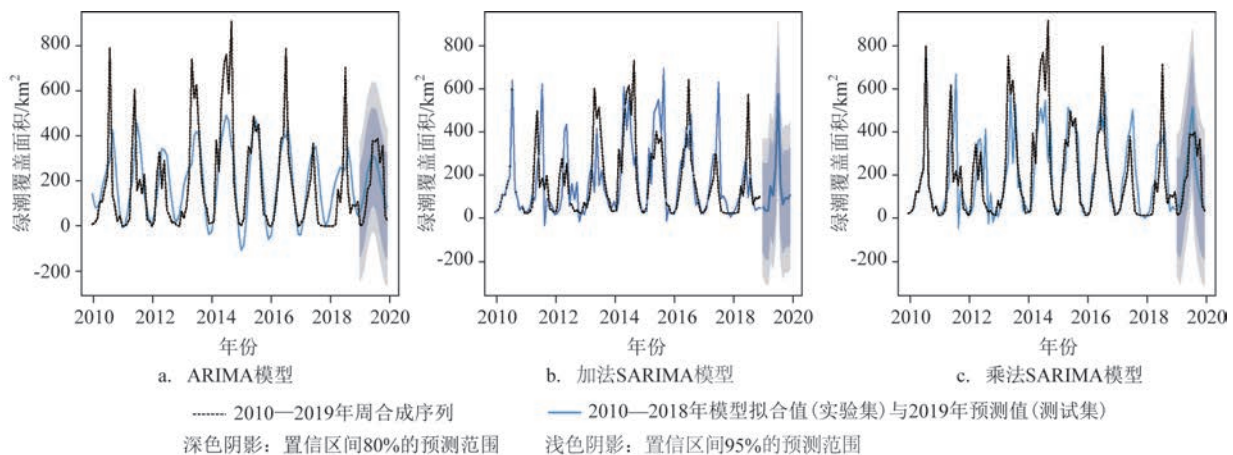


图7 2010—2018年绿潮覆盖面积的模型模拟及2019年预测结果

Fig.7 Green tide coverage area of model simulation from 2010—2018 and prediction results in 2019

3 结论与展望

本文基于3种不同时间序列模型尝试建立绿潮覆盖面积的时间序列预测方法,并对不同模型的预测效果做了对比。结果表明:

①通过对2010—2019年5月8日—8月7日MODIS成像条件较好的绿潮覆盖面积数据进行周平均处理后,建立了绿潮覆盖面积周合成等时间间隔序列。根据统计,遥感监测的绿潮出现时间主要集中于5月中下旬,峰值集中于6月中下旬,消亡时间集中于7月下旬—8月上旬。遥感监测表明绿潮每年平均持续时间为84 d,峰值面积平均为576 km²。

②将原始序列分为训练集(2010—2018年)和测试集(2019年),通过平稳性检验建立原始序列非季节性模型ARIMA(2,0,2),并通过一阶12步差分建立加法季节性模型SARIMA(1,0,0)×(0,1,0)₁₂和乘法季节性模型SARIMA(1,0,0)×(0,1,1)₁₂,3个模型都通过模型白噪声检验和参数检验,说明都具有较好的模拟效果。从模拟效果看,乘法SARIMA模型的AIC值最小,为1 269.02,其次为加法SARIMA模型,AIC=1 297.5,ARIMA(2,0,2)的AIC值最大,为1 409.76,说明乘法SARIMA模型的模拟效果最优。对2019年绿潮覆盖面积进行预测,乘法季节性模型SARIMA(1,0,0)×(0,1,1)₁₂的MAE最小,为95.56 km²,ARIMA(2,0,2)的RMSE最小,为154.87 km²,SARIMA(1,0,0)×(0,1,0)₁₂的预测精度最低。季节性模型预测的峰值出现时间与实测值更为接近。

由于时间序列预测法基于数据驱动,对机理研究的要求较低,因此在影响因素较为复杂的绿潮生消机制尚未清楚的情况下,可以通过提取长期趋势、季节波动和不规则成分,实现绿潮年度预测应用。该方法简单实用,可作为绿潮年度趋势预测方法的有效补充。在今后的相关研究中,还需结合生态动力学模型弥补数据预处理的不足,加强将绿潮相关影响因素纳入时间序列模型的研究,探索数据分解与平滑技术与ARMA模型相结合的混合模型预测方法,以期为绿潮长期预测提供技术支撑。

参考文献:

[1] 国家海洋局. HY/T 217—2017 绿潮预报和警报发布[S]. 北京: 中

国标准出版社, 2017.

State Oceanic Administration. HY/T 217—2017 The issue of green tide forecasts and warnings[S]. Beijing: Standards Press of China, 2017.

[2] 于仁成, 吕颂辉, 齐雨藻, 等. 中国近海有害藻华研究现状与展望[J]. 海洋与湖沼, 2020, 51(4): 768-788.

YU R C, LYU S H, QI Y Z, et al. Progress and perspectives of harmful algal bloom studies in China[J]. Oceanologia et Limnologia Sinica, 2020, 51(4): 768-788.

[3] YE N H, ZHANG X W, MAO Y Z, et al. 'Green tides' are overwhelming the coastline of our blue planet: taking the world's largest example[J]. Ecological Research, 2011, 26(3): 477-485.

[4] WANG Z L, XIAO J, FAN S L, et al. Who made the world's largest green tide in China? - an integrated study on the initiation and early development of the green tide in Yellow Sea[J]. Limnology and Oceanography, 2015, 60(4): 1105-1117.

[5] 高姗, 刘桂梅, 王辉, 等. 海洋生态预报的研究进展与发展应用[J]. 海洋预报, 2020, 37(2): 84-93.

GAO S, LIU G M, WANG H, et al. Research progress and developmental application of marine ecological forecasting[J]. Marine Forecasts, 2020, 37(2): 84-93.

[6] GENG H X, YU R C, ZHANG Q C, et al. Tracing the settlement region of massive floating green algae in the Yellow Sea[J]. Journal of Oceanology and Limnology, 2019, 37(5): 1555-1565.

[7] CUI J J, ZHANG J H, MONOTILLA A P, et al. Assessment of blooming Ulva macroalgae production potential in the Yellow Sea, China[J]. Phycologia, 2019, 58(5): 535-541.

[8] GALLARDO-RODRÍGUEZ J, ASTUYA-VILLALÓN A, LLANOS-RIVERA A, et al. A critical review on control methods for harmful algal blooms[J]. Reviews in Aquaculture, 2019, 11(3): 661-684.

[9] GAO Y C, DONG Y H, LI H T, et al. Influence of environmental factors on spatial-temporal distribution patterns of dinoflagellate cyst communities in the South China Sea[J]. Marine Biodiversity, 2019, 49(2): 769-781.

[10] 李成鹏. 基于绿潮空间分布形态和生存环境的生物量计算及预测[D]. 青岛: 山东科技大学, 2020.

LI C P. Biomass calculation and prediction based on green tide spatial distribution and living environment[D]. Qingdao: Shandong University of Science and Technology, 2020.

[11] 刘赞明. 基于大数据技术的绿潮预测研究[D]. 青岛: 青岛科技大学, 2020.

LIU Y Y. Research on green tide forerast based on big data technology[D]. Qingdao: Qingdao University of Science and Technology, 2020.

[12] 乔方利, 王关锁, 吕新刚, 等. 2008与2010年黄海浒苔漂移输运特征对比[J]. 科学通报, 2011, 56(18): 1470-1476.

QIAO F L, WANG G S, LYU X G, et al. Drift characteristics of green macroalgae in the Yellow Sea in 2008 and 2010[J]. Chinese Science Bulletin, 2011, 56(21): 2236-2242.

- [13] 白涛, 黄娟, 高松, 等. 黄海绿潮应急预测系统业务化研究与应用[J]. 海洋预报, 2013, 30(1): 51-58.
BAI T, HUANG J, GAO S, et al. Operational forecast system for green tides in the Yellow Sea[J]. Marine Forecasts, 2013, 30(1): 51-58.
- [14] 李燕, 李云, 刘钦政. 浒苔漂移轨迹预报系统[J]. 海洋预报, 2010, 27(4): 74-78.
LI Y, LI Y, LIU Q Z. The operational enteromorpha drifting forecasting system[J]. Marine Forecasts, 2010, 27(4): 74-78.
- [15] 何恩业, 季轩梁, 高姗, 等. 黄海浒苔漂移输运和生长消亡过程的数值模拟与预测应用[J]. 海洋与湖沼, 2021, 52(1): 39-50.
HE E Y, JI X L, GAO S, et al. Numerical simulation and forecasting of drift, growth, and death of Enteromorpha in the Yellow Sea[J]. Oceanologia et Limnologia Sinica, 2021, 52(1): 39-50.
- [16] 刘旭, 田杰, 吴彬锋, 等. 基于MODIS的2019年中国南黄海绿潮生消过程分析[J]. 海洋预报, 2021, 38(5): 53-59.
LIU X, TIAN J, WU B F, et al. Analysis of the formation and decay progress of the green tide in the southern Yellow Sea in 2019 based on MODIS[J]. Marine Forecasts, 2021, 38(5): 53-59.
- [17] 唐莹莉. 基于支持向量机的绿潮灾害发展过程中影响因素的权重分析[D]. 上海: 上海海洋大学, 2016.
TANG Y L. Weight analysis of each influence factor during the green tide disaster based on SVM[D]. Shanghai: Shanghai Ocean University, 2016.
- [18] 何世钧, 周媛媛, 张婷, 等. 基于主导因子的绿潮灾害预测方法研究[J]. 海洋环境科学, 2018, 37(3): 326-331.
HE S J, ZHOU Y Y, ZHANG T, et al. Development of a preliminary model for forecasting development stages of "Green tide" disaster based on analyses of dominant environmental factors[J]. Marine Environmental Science, 2018, 37(3): 326-331.
- [19] 孙丰霖. 中国沿海赤潮灾害时间序列特征的模拟与预测[J]. 海洋通报, 2021, 40(2): 232-240.
SUN F L. Simulation and forecast of the red tide's time series characteristics in China seas[J]. Marine Science Bulletin, 2021, 40(2): 232-240.
- [20] 余璇, 石绥祥, 徐凌云, 等. 基于深度学习的赤潮发生预报方法研究[J]. 海洋通报, 2021, 40(5): 566-577.
YU X, SHI S X, XU L Y, et al. Research on red tide occurrence forecast based on deep learning[J]. Marine Science Bulletin, 2021, 40(5): 566-577.
- [21] 徐丽丽, 余璇, 高鑫鑫, 等. 一种基于时间序列分析的赤潮预测新方法研究——以浙江海域为例[J]. 海洋预报, 2020, 37(5): 95-103.
XU L L, YU J, GAO X X, et al. Study on the red tide prediction based on time series analysis——A case study in Zhejiang sea area[J]. Marine Forecasts, 2020, 37(5): 95-103.
- [22] 徐海龙, 谷德贤, 张文亮, 等. 基于时间序列的海洋赤潮灾害特征分析[J]. 海洋通报, 2014, 33(4): 469-474.
XU H L, GU D X, ZHANG W L, et al. Analysis of the red tide features based on time series in the China Sea[J]. Marine Science Bulletin, 2014, 33(4): 469-474.
- [23] 史代敏, 谢小燕. 应用时间序列分析[M]. 2版. 北京: 高等教育出版社, 2019: 166-209.
SHI D M, XIE X Y. Application of time series analysis[M]. 2nd ed. Beijing: Higher Education Press, 2019: 166-209.

Simulation and prediction of the characteristics of green tide disaster time series in China

LIU Xu^{1,2,3}, JIANG Shan¹, WANG Zheng¹, LIANG Yingqi¹, HE Enye^{1*}

(1. National Marine Environmental Forecasting Center, Beijing 100081 China; 2. School of Economics & Management, Beijing Forestry University, Beijing 100083, China; 3. Key Laboratory of Marine Hazards Forecasting, National Marine Environmental Forecasting Center, Ministry of Natural Resources, Beijing 100081, China)

Abstract: Based on the satellite remote sensing images of green tide in the Yellow Sea from 2010 to 2019, an analysis method of the green tide coverage area time series is constructed in this paper. The remote sensing monitoring data with good imaging conditions from May 8th to August 7th each year is preprocessed into a weekly average time serie. The years from 2010 to 2018 are set as the model training set, and the year 2019 is set as the model validation set. Based on the Box-Jenkins method, the Autoregressive Integrated Moving Average ARIMA (2,0,2), additive seasonal model SARIMA (1,0,0) \times (0,1,0)₁₂ and multiplicative seasonal model SARIMA (1,0,0) \times (0,1,1)₁₂ are constructed, which all pass the model white noise test and parameter significance test with good simulation effect and predictability. Specially, the Akaike Information Criterion (AIC) value of SARIMA

$(1,0,0) \times (0,1,1)_{12}$ is the smallest with the mean absolute error (MAE) and root mean squared error (EMSE) of 95.56 km^2 and 156.74 km^2 , respectively in 2019, which improved Compared with ARMA (2,0,2), the MAE increases by 12% and the RMSE decreases by 1.2%. The prediction accuracy of SARIMA $(1,0,0) \times (0,1,0)_{12}$ is the lowest with the MAE and RMSE of 115.12 km^2 and 192.16 km^2 , respectively.

Key words: green tide; ARIMA; SARIMA; time series; remote sensing monitor