

# 基于 PLS 和组合预测方法的冬小麦收获 指数高光谱估测\*

陈 帼<sup>1, 2, 3</sup>, 徐新刚<sup>2, 3\*</sup>, 杜晓初<sup>1</sup>, 杨贵军<sup>2, 3</sup>, 赵晓庆<sup>2, 3</sup>,  
魏鹏飞<sup>1, 2, 3</sup>, 王玉龙<sup>2, 3</sup>, 范玲玲<sup>2, 3</sup>

(1. 湖北大学资源与环境学院, 武汉 430062; 2. 农业部农业遥感机理与定量遥感重点实验室/北京农业信息技术研究中心, 北京 100097; 3. 国家农业信息化工程技术研究中心, 北京 100097)

**摘要:**【目的】通过遥感反演测量收获指数(HI),可节省时间和人力,但需要提高精度。通过权重最优组合算法改善收获指数估算精度,为基于多时相多光谱信息的HI遥感估算提供新方法参考。【方法】利用测定的冬小麦多个关键生育期的冠层光谱数据,对筛选的44种常用植被指数与实测收获指数进行相关性分析,挑选出每个生育期中5种最优的典型植被指数;应用偏最小二乘(PLS)的方法建模,分别得到基于单个生育期光谱信息的HI遥感估测模型;借鉴组合预测原理,应用组合预测方法对全部单生育期的各HI光谱模型赋予最优权重,最终构建基于多生育期数据的HI光谱组合预测模型。【结果】(1)利用PLS后,单一生育期的建模结果较单一植被指数有所改进,但仍有待提高;(2)应用组合预测原理的HI组合预测模型,显著改善了HI的估测精度, $R^2$ 达到0.55,较单生育期的建模预测,提升了13%。【结论】基于多生育期信息的组合预测方法,对各单一生育期HI预测模型赋予最优权重进行优化组合,实质间接利用了各生育期对作物HI形成的贡献,显著提高冬小麦收获指数的估测精度,是一种新颖的作物HI遥感估测方法。

**关键词:** 收获指数; 偏最小二乘法; 冬小麦; 组合预测法; 遥感光谱; 多生育期

DOI: 10.12105/j.issn.1672-0423.20190204

## 0 引言

收获指数(Harvest Index, HI)指收获时作物籽粒产量和地上部生物量的比值,又名经济系数,是选择作物品种和品种改良研究中的重要参考因子<sup>[1]</sup>。收获指数通常可以通过田间取样获得,但需要耗费较大的人力、物力,并具有一定滞后性。遥感技术拥有动态、快速和准确获取地表作物参数信息的优势,随着遥感监测技术的快速发展,通过遥感手段获取收获指数将成为未来收获指数评价的发展趋势<sup>[2]</sup>。

当前,基于遥感技术的作物收获指数遥感估算已开展了初步研究。杜鑫等<sup>[3]</sup>利用

收稿日期: 2019-03-05

第一作者简介: 陈帼(1990—), 硕士研究生。研究方向: 农业遥感。Email: micg0421@163.com

※ 通信作者简介: 徐新刚(1976—), 博士、研究生导师。研究方向: 农业遥感应用基础。Email: xxgpaper@126.com

\* 基金项目: 国家自然科学基金项目(41571416); 国家重点研发计划项目(2017YFD0201501)



2019年4月

遥感技术监测作物收获指数的可行性分析研究, 基于 HI 的形成过程总结归纳出 3 类利用遥感技术实现 HI 估测的方法, 为如何利用遥感方式估算收获指数提供了借鉴思路。Moriondo 等<sup>[4]</sup> 基于时间序列 NDVI (归一化植被指数) 遥感信息, 利用小麦开花前后 NDVI 均值比作为指示特征, 开展小麦的收获指数遥感提取研究, 但是该方法需要首先确定提取区域内最大收获指数和其可能会有的变动幅度, 不同的取值一定程度会影响收获指数估算结果。基于 Moriondo 等的研究, 任建强等<sup>[5]</sup> 做了方法改进, 利用 MODIS 卫星遥感影像获取冬小麦生长期时间序列 NDVI 数据, 以小麦开花前的 NDVI 累计值与花后 NDVI 累计值的比值作为指示 HI 的遥感指数, 很好地估测了区域冬小麦的收获指数, 反演结果与实测结果的  $R^2$  达到了 0.49。上述方法, 尽管很好地实现区域尺度作物 HI 的遥感估测, 但需要作物生长季内逐日或者短周期间隔的时间序列卫星遥感影像, 数据处理工作量巨大。

另一方面, 由于作物收获指数反映了作物光合产物在籽粒和营养器官上的分配比例, 它与作物不同生育期植株体内光合产物的形成、运转过程息息相关, 并受到诸多因素的影响, 因此, 作物不同生育期的生长对作物最终收获指数的形成都具有或多或少的的影响, 如何来刻画不同生育期对作物 HI 的影响贡献, 组合预测方法提供了一种可能途径。文章通过获取的冬小麦多生育期多时相地面冠层高光谱数据, 开展小区田块尺度的作物收获指数高光谱估测研究, 通过引入最优组合预测方法, 将基于不同生育期光谱信息建立的不同冬小麦收获指数光谱估测模型进行组合, 通过优化算法赋予最优权重, 从而构建组合估测模型实现冬小麦收获指数的高光谱估测, 相关的研究并不多见。该文提出的作物收获指数遥感光谱估测方法, 尝试利用权重最优组合算法, 以达到充分利用作物多个生育期有用信息进而改善收获指数估算精度的目的, 以期为基于多时相光谱信息的 HI 遥感估算提供新的方法参考。

## 1 研究区与研究方法

### 1.1 研究区域

实验区位于北京郊区的顺义区和通州区, 该区域属于暖温带半湿润大陆性季风气候。年平均温度 11.3℃, 年平均降水量在 620 mm 左右。气候宜人, 地形平缓, 适宜耕种。为展开研究, 在顺义区和通州区种植基地中随机挑选出 27 个采样点, 采样选取大田块、生长具有代表性冬小麦地块进行实验获取数据。

### 1.2 冬小麦冠层光谱测定

实验于 2009—2010 年冬小麦生长季开展, 分别选取冬小麦起身期、拔节期、孕穗期、开花期和灌浆期, 对冬小麦冠层进行光谱测定。测定日期分别为: 2010 年 4 月 1 日、4 月 17 日、4 月 29 日、5 月 17 日和 6 月 2 日。测量光谱仪选用 ASD Field Spec FR2500, 光谱范围 350~2 500 nm, 采样间隔为 1 nm。尽量选择在天气晴朗时测定, 测定时间为北京时间 10: 00~14: 00。测定时, 探头始终垂直向下且保持与地面约为 1 m 的距离, 探头视场角为 25°。在每个采样点测量 10 次, 取平均值为光谱测定的结果。在每个采样点测

定前、后立即进行标准白板矫正。在进行植被指数计算时，为排除干扰波段，仅仅选用敏感度较高的 400~1 100 nm 作为有效分析数据。

### 1.3 收获指数测定

选取研究区内共 27 个实测样区。其中，通州区、顺义区样区数分别为 15 个和 12 个。样区位置的选择充分考虑了冬小麦生长状况和区域分布的代表性，且样区面积均不小于 100 m × 100 m。采用五点取样法进行取样，拷种称量后，剪掉根部，只保留地上生物量部分进行晾晒，之后称量 5 采样点冬小麦干重记录为  $M$ 。随后进行脱粒处理，称取籽粒的重量记录为  $M'$ ，这样每个采样点的实测收获指数 HI 即为两者重量的比值。该实验测得的结果 HI，即为试验区域内冬小麦收获指数的实际测量值。该值将被用于构建单个生育期的预测模型以及验证组合预测模型所计算结果的精度。计算公式为：

$$HI = \frac{M'}{M} \quad (1)$$

式 (1) 中， $HI$  为收获指数， $M'$  为小麦籽粒重量， $M$  为小麦干重。

### 1.4 植被指数选择

到目前为止，植被指数的种类已经达到了 100 多种，并在各个领域得到了广泛应用。由于关于收获指数的研究成果有限，能够明显反映收获指数大小的植被指数还在探索中，但是收获指数的大小与作物光合作用关系密切<sup>[6]</sup>，在选择植被指数构建模型时，可以将对光合作用有关的因素列入考虑范围，如叶面积大小、温度、施氮水平、叶绿素含量等。

在挑选出的植被指数中，NDVI（归一化差值植被指数）、DVI（差值植被指数）、RVI（比值植被指数）、EVI（增强植被指数）、OSAVI（优化土壤调整植被指数）等均属于常用的植被指数，而后学者们为了估测水稻大叶面积指数，在 NDVI 的计算公式上进行了改良，提出了 GBNDVI（绿蓝波段归一化植被指数）、GRNDVI（红绿波段归一化植被指数）、RBNDVI（红蓝波段归一化植被指数）、BNDVI（蓝波段归一化植被指数），这些改良的植被指数也被挑选为反演参数。常用于监测叶绿素含量的植被指数有：NPCI（归一化色素叶绿素植被指数）、TCARI（转化叶绿素吸收反射指数）、MCARI 2#（改进叶绿素吸收比率指数 II）、GVI（绿度植被指数）等；植被含氮量估测常用的植被指数有：NDRE（归一化红边指数）、MTCI（MERIS 陆地叶绿素指数）、SIPI（结构不敏感色素指数）、PPR（植被色素比率）等；除此之外，该研究还引入了一些新型的植被指数作为对比参考，如 DPI（双峰值指数）、REP-li（线性内插法红边位置指数）、PSRI（三波段比值指数）、MSR（改进比值植被指数）等。

综合以上，共计 44 种植被指数被列为反演参数。计算时，分别选择 620~760 nm、492~577 nm、400~450 nm、700~1 100 nm 的波段范围内平均反射率作为红光（R）、绿光（G）、蓝光（B）和近红外（NIR）波段的反射率值。其详细的计算公式和引用文献如表 1 所示。

2019年4月

表 1 采用的植被指数  
Table 1 Summary of vegetation indices studied

植被指数	名称	公式	参考文献
NDVI	归一化植被指数	$(NIR-R) / (NIR+R)$	[ 7 ]
RVI	比值植被指数	$NIR/R$	[ 8 ]
DVI	差值植被指数	$NIR-R$	[ 9 ]
WI I <sup>#</sup>	水分植被指数	$R_{900}/R_{970}$	[ 10 ]
EVI	增强植被指数	$2.5 (NIR-R) / (NIR+6 \times R-7.5 \times B+1.5)$	[ 11 ]
TCARI	转化叶绿素吸收反射指数	$3 ((R_{700}-R_{670}) -0.2 (R_{700}-R_{550}) (R_{700}/R_{670}))$	[ 12 ]
OSAVI	优化土壤调节植被指数	$1.16 (R_{800}-R_{670}) / (R_{800}+R_{670}+0.16)$	[ 12 ]
TCARI/OSAVI	整合指数 1	TCARI/OSAVI	[ 12 ]
MTCI	MERIS 陆地叶绿素指数	$(R_{750}-R_{710}) / (R_{710}-R_{680})$	[ 12 ]
OSAVI × CI <sub>red edge</sub>	整合指数 2	OSAVI × CI <sub>red edge</sub>	[ 12 ]
RENDVI	红边归一化植被指数	$(R_{750}-R_{710}) / (R_{710}-R_{680})$	[ 12 ]
CI <sub>red edge</sub>	红边模型	$R_{750}/R_{720}-1$	[ 13 ]
DCNI I	双峰冠层氮指数	$(R_{720}-R_{700}) / (R_{700}-R_{670}) / (R_{720}-R_{670}+0.03)$	[ 14 ]
WDRVI	宽范围动态植被指数	$(0.1 \times NIR-NIR) / (0.1 \times NIR+R)$	[ 15 ]
MSAVI	修改型土壤调整植被指数	$2 \times NIR+1 - (\sqrt{2 (NIR+1)^2 - 8 (NIR-R) / 2})$	[ 16 ]
PPR	纸杯色素比率	$(R_{550}-R_{450}) / (R_{550}+R_{450})$	[ 17 ]
SIPI	结构不敏感色素指数	$(R_{800}-R_{445}) / (R_{800}-R_{680})$	[ 18 ]
MTVI 2 <sup>#</sup>	改善角度植被指数 2 <sup>#</sup>	$1.5 \times (1.2 \times (R_{800}-R_{550}) - 2.5 \times (R_{670}-R_{550}))$	[ 19 ]
SAVI	土壤调节植被指数	$1.5 \times (NIR-R) / (NIR+R+0.5)$	[ 20 ]
NDGI	归一化差异绿度指数	$(G-R) / (G+R)$	[ 21 ]
RI	红色植被指数	$(R-G) / (R+G)$	[ 22 ]
NDRE	标准化差分红边植被指数	$(R_{790}-R_{720}) / (R_{790}+R_{720})$	[ 23 ]
RRWVI	红边抗水植被指数	$R_{970} \times (R_{790}-R_{720}) / R_{900} \times (R_{790}+R_{720})$	[ 24 ]
NVI	新植被指数	$(R_{777}-R_{747}) / R_{673}$	[ 24 ]
NPCI	归一化色素叶绿素植被指数	$(R_{680}-R_{430}) / (R_{680}+R_{430})$	[ 25 ]
PRI	光辐射指数	$(R_{531}-R_{570}) / (R_{531}+R_{570})$	[ 26 ]
MSR	改进比值植被指数	$(NIR/R-1) / (\sqrt{NIR/R} + 1)$	[ 27 ]
GRNDVI	红绿波段归一化植被指数	$(R-G) / (R+G)$	[ 28 ]

续表

植被指数	名称	公式	参考文献
GBNDVI	绿蓝波段归一化植被指数	$(\text{NIR}-\text{G}-\text{B}) / (\text{NIR}+\text{G}+\text{B})$	[29]
RBNDVI	红蓝波段归一化植被指数	$(\text{NIR}-\text{R}-\text{B}) / (\text{NIR}+\text{R}+\text{B})$	[29]
BNDVI	蓝波段归一化植被指数	$(\text{NIR}-\text{B}) / (\text{NIR}+\text{B})$	[29]
GNDVI	绿波段归一化植被指数	$(\text{NIR}-\text{G}) / (\text{NIR}+\text{G})$	[29]
RDVI	再归一化植被指数	$(\text{NIR}-\text{R}) / \sqrt{\text{NIR}+\text{R}}$	[30]
RGRI	红边比值指数	R/G	[31]
TNDVI	转换型植被指数	$\sqrt{\frac{(\text{NIR}-\text{R})}{\text{NIR}+\text{R}}+0.5}$	[31]
GVI	绿度植被指数	NIR/G	[31]
MCARI 2 <sup>#</sup>	改进叶绿素吸收比率指数 II	$\frac{1.5 \times (2.5 \times (\text{R}_{800}-\text{R}_{670}) - 1.3 \times (\text{R}_{800}-\text{R}_{550}))}{\sqrt{(2\text{R}_{800}+1)^2 - (6\text{R}_{800}-5\sqrt{\text{R}_{670}}) - 0.5}}$	[32]
REP-li	线性内插法红边位置	$700+40 \left( \frac{(\text{R}_{670}+\text{R}_{780})}{2}-\text{R}_{700} \right) / (\text{R}_{740}-\text{R}_{700})$	[33]
$(\text{R}_{683})^2 / (\text{R}_{655} \times \text{R}_{691})$	曲率光学指数	$(\text{R}_{683})^2 / (\text{R}_{655} \times \text{R}_{691})$	[33]
PSRI	三波段比值指数	$(\text{R}_{678}-\text{R}_{500}) / \text{R}_{750}$	[33]
DPI	双峰值指数	$(\text{R}_{688}+\text{R}_{710}) / (\text{R}_{697})^2$	[33]
FRI690/600	Fluorescence Ratio Index 690/600	$\text{R}_{690} / \text{R}_{600}$	[33]
FRI740/800	Fluorescence Ratio Index 740/800	$\text{R}_{740} / \text{R}_{800}$	[33]
mND705	改进归一化植被指数 705	$(\text{R}_{750}-\text{R}_{705}) / (\text{R}_{750}+\text{R}_{705}-2\text{R}_{445})$	[33]

## 1.5 统计分析

将测定出的冬小麦 5 个生育期的冠层光谱数据，代入表 1 所列公式进行计算。按照不同生育期，排列出 27 个取样点对应的 44 个植被指数值，并将每个生育期的 44 种植被指数与冬小麦实测收获指数 HI 之间建模，进行相关性计算。选择决定性系数 ( $R^2$ ) 和均方根误差 (RMSE) 作为指标，将所计算的结果进行排序，挑选  $R^2$  较大同时 RMSE 尽量较小的 5 个植被指数作为筛选结果。

## 1.6 偏最小二乘建模

偏最小二乘法 (Partial Least-Square, PLS) 是一种多元统计分析法，具有计算简单、建模预测精度高、解释性强等优点。PLS 通常用于解决对于多个因变量和多个自变量的回归建模等问题，当解释变量的个数远超过样本个数或者解释变量内部存在多重共线性时，PLS 能运用成分提取的方法，解释变量与因变量的相关性<sup>[34]</sup>。研究中筛选出的植被指数即为解释变量，实测收获是因变量，将二者利用 PLS 结合后，得到每个生育

2019年4月

期的单个预测方程。与其他线性模型一样，PLS的最终结果也是一个线性模型，其方程为：

$$Y=A_1X_1+A_2X_2+\cdots+A_nX_n+B \quad (n=1, 2, \cdots, N) \quad (2)$$

式(2)中，Y是因变量，即实测收获指数； $X_1 \sim X_n$ 是用于构建模型的植被指数； $A_1 \sim A_n$ 是对应植被指数的系数；B是残差参数。使用PLS得到 $A_1 \sim A_n$ 和B的值，即可得到单个生育期的预测方程。

### 1.7 组合预测法

组合预测法是采用两种或者两种以上不同的预测方法，对同一对象进行预测，对各单独的预测结果适当加权综合后作为其最终结果<sup>[35]</sup>。这种方法能聚集各个预测方法的有用信息，使得组合模型的精度优于其中任意一个单一模型的模拟精度，从而达到提高预测收获指数精度的目的。研究采用组合预测法确定每个生育期对于最终收获指数值的贡献程度，即计算出每个生育期对应的权重。确定权重的算法很多，该文采用最优加权算法，评价最优的标准为使组合预测偏差值之和最小。公式为：

$$f_i = \sum_{i=1}^n k_i f_{it} \quad (t=1, 2, \cdots, N; i=1, 2, \cdots, N) \quad (3)$$

式(3)中， $f_i$ 代表第*i*种预测方法在*t*时刻的预测值，为第*i*种预测方法的权系数，且满足

$$k_i \geq 0, \sum_i k_i = 1$$

根据最小二乘法，当残差平方和达到最小，变权系数达到最佳。

设 $e_t$ 为*t*时刻组合预测的偏差，则计算最佳变权系数的方程组为：

$$\begin{cases} k_1+k_2+\cdots+k_N=1 \\ e_1k_1+e_2k_2+\cdots+e_Nk_N-m_t+n_t=0 \\ k_i \geq 0 \quad (i=1, 2, \cdots, n) \\ m_t \geq 0, n_t \geq 0 \quad (t=1, 2, \cdots, n) \\ \min S = \sum_{t=1}^n (m_t+n_t) \end{cases} \quad (4)$$

式(4)中， $k_1 \sim k_N$ 是需要求解的系数，代表单个生育期的预测方程在组合预测中的权重。最后，用实测的数据和组合方程预测出的收获指数值进行验证。

## 2 研究结果

### 2.1 相关性分析

将5个生育期内的植被指数和实测植被指数进行相关性分析，计算出 $R^2$ 和RMSE，并进行排序，选取每个时期中 $R^2$ 尽可能大和RMSE尽可能小的5个最优植被指数，结果如表2所示。每一项植被指数与收获指数的 $R^2$ 都非常低，均没有超过0.3，拔节期内所有植被指数与实测收获指数的相关性几乎趋近于0，且RMSE也出现了非常巨大的值。可以看出，单个生育期的单个植被指数与实测收获指数的相关性太低，因此筛选5个最优植被指数与收获指数构建模型。

表 2 不同生育期最优植被指数  
Table 2 Optimal vegetation indices in different stages

生育期	植被指数	R <sup>2</sup>	RMSE
起身期	FRI 690/600	0.222	0.669
	WI 1 <sup>#</sup>	0.203	0.587
	PRI	0.196	0.340
	PSRI	0.194	0.322
	DPI	0.194	0.185
拔节期	TVI	0.003	1 085.356
	REP-li	0.001	724.655
	MTVI 2 <sup>#</sup>	0.004	74.210
	DVI	0.001	26.231
	$(R_{683})^2 / (R_{655} \times R_{691})$	0.115	22.560
孕穗期	TCARI/OSAVI	0.142	14.300
	TCARI	0.140	8.749
	DCNI I	0.136	0.178
	MTVI 2 <sup>#</sup>	0.080	82.098
	DVI	0.057	28.569
开花期	DCNI I	0.198	0.130
	MTCI	0.179	4.262
	TCARI	0.161	6.211
	TCARI/OSAVI	0.155	10.230
	REP-li	0.124	727.516
灌浆期	MTCI	0.223	1.964
	RVI	0.171	2.103
	OSAVI $\times$ CI <sub>red edge</sub>	0.168	0.266
	RRWVI	0.168	0.545
	DCNI I	0.168	0.220

## 2.2 单生育期与收获指数的 PLS 建模

利用 5 个生育期中的最优植被指数与 PLS 结合，建立冬小麦收获指数的估算模型，可以得到经过建模后，单生育期的预测模型。将模型中得到的预测值与实测的收获指数进行相关性分析，按照生育期进行顺序排列，结果如表 3 所示。相对于单个植被指数拟合，单生育期的多个植被指数拟合的结果在相关性方面有了明显提高，能明显看出单个生育期预测方程对于最终结果的影响大小。但是，只有灌浆期拟合结果的 R<sup>2</sup> 超过了 0.4，

2019年4月

其余4个时期与实测值的相关性都不明显；5个RMSE相较于单个生育期拟合而言，结果有了明显优化，且结果相对稳定，没有太大起伏。说明在经过PLS建模之后，单个生育期的拟合已达到最优状态，但是精度并不足以支持预测方程成立，需要进一步提高。

表3 植被指数与PLS结合预测收获指数的结果

Table 3 Performance of the combination of vegetation indices and PLS for predicting harvest index

生育期	R <sup>2</sup>	RMSE
起身期	0.213	0.069
拔节期	0.381	0.062
孕穗期	0.359	0.063
开花期	0.329	0.064
灌浆期	0.423	0.059

注:  $r(0.05, 27) = 0.381$ ;  $r(0.01, 27) = 0.487$

### 2.3 组合预测建模

利用组合预测的方法，计算出了最优变权系数，将系数得到组合预测的方程为：

$$y = 0.043f_{\text{起身}} + 0.164f_{\text{拔节}} + 0.038f_{\text{孕穗}} + 0.53f_{\text{开花}} + 0.225f_{\text{灌浆}} \quad (5)$$

式(5)中每个单生育期预测方程前的系数，代表单个生育期对最终结果的贡献程度，即在组合预测方程中，单生育期预测方程的权重系数。这些系数根据前文中式(4)所列出的方程计算而出。

将各生育期的预测结果带入组合预测方程中，得到最终的预测结果。建立实测地面冬小麦收获指数与组合预测结果的关系为：

$$y = 1.5554x - 0.2254 \quad (R^2 = 0.553) \quad (6)$$

式(6)中， $x$ 为组合预测得到的参数， $y$ 为预测出的收获指数，详细结果由图1所示。由图可以看出，相较于单个生育期的预测结果而言，RMSE的变化不大，比单个预测模型中的最小值只降低了0.003，但是 $R^2$ 有了显著提升，较单个预测模型的最大值提升了0.13。根据前文所述，组合预测法具有提取单个预测模型中有效信息的功能，单个预测模型的有效信息越多，代表单个模型的贡献度越大，在组合预测中被赋予的权重就越大，反之则越小。由式(5)可以看出，开花期和灌浆期的单个预测模型在组合中权重最大，这与冬小麦栽培种植的特征基本吻合。由于前3个生育期的时间较短，作物光合产物的累积量不及后两个生育期的累积量充足，而收获指数的确定与光合作用累计的生物量密切相关，所以时间间隔最长的开花期相较于其他生育期对于收获指数的影响力最大，因此相关性最高，权重最大。因此，运用组合预测的算法，可以成功筛选出单个预测模型中的有效信息，利用最优加权组合预测模型的精度比起单个模型预测的精度提升明显。表明组合预测的方法能够为利用遥感数据信息估算收获指数提供有效的方法途径。

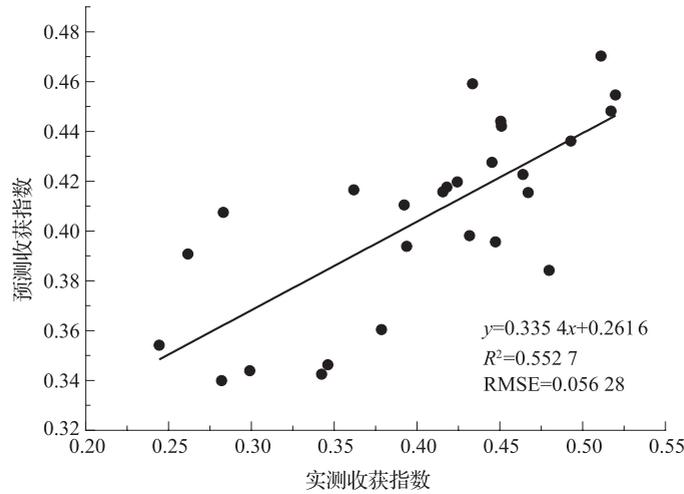


图 1 组合预测值与冬小麦收获指数关系

Fig.1 Relationship between combination predicted value and winter wheat harvest index

### 3 结论

单一生育期的植被指数与实测收获指数相关性较低，在挑选出每个生育期中相关性最合适的 5 个植被指数后与 PLS 建模，相关性有所提高但依然很低，值最高的  $R^2$  来自灌浆期为 0.42，对应的 RMSE 为各生育期预测值中最低为 0.06，说明 PLS 的确具有能提升预测精度的能力，可是在该研究中，PLS 与单一生育期建模并不能达到理想的预测标准。在利用组合预测法构建方程以后，预测值与实测值的相关性有了明显提高。研究结果表明，在多种不同的方法进行预测时，使用组合预测法可以有效起到提取有用信息的作用，从而可以提高对冬小麦收获指数的估算精度。利用组合预测法得到的预测结果与实测结果的  $R^2$  能达到 0.55，对应的 RMSE 为 0.06。

### 参考文献

- [ 1 ] 张福春, 朱志辉. 中国作物的收获指数. 中国农业科学, 1990, 23(2): 83~87.
- [ 2 ] 潘晓华, 邓强辉. 作物收获指数的研究进展. 江西农业大学学报, 2007, 29(1): 1~5.
- [ 3 ] 杜鑫, 吴炳芳, 蒙继华, 等. 基于遥感技术监测作物收获指数 (HI) 的可行性分析. 中国农业气象, 2010, 31(3): 453~457.
- [ 4 ] Moriondo M, Maselli F, Bindi M. A simple model of regional wheat yield based on NDVI data. *European Journal of Agronomy*, 2007, 26(3): 266~274.
- [ 5 ] 任建强, 陈仲新. 基于时序归一化植被指数的冬小麦收获指数空间信息提取. 农业工程学报, 2010, 26(8): 160~167.
- [ 6 ] 潘晓华, 邓强辉. 作物收获指数的研究进展. 江西农业大学学报, 2007, 29(1): 1~5.
- [ 7 ] Rouse J W, Hass R H, Deering D W, et al. Monitoring the vernal advancement and retrogradation (Green Wave Effect) of natural vegetation. *NASA/GSFC Type III Final Report, Greenbelt, MD. US: NASA*, 1974: 371.
- [ 8 ] Birth G S, Mcvey G. Measuring the color of growing turf with a reflectance spectrophotometer. *Agronomy Journal*, 1968, 60: 640~643.
- [ 9 ] Jordan C F. Derivation of leaf area index from quality of light on the forest floor. *Ecology*, 1969, 50: 663~666.



2019年4月

- [ 10 ] Penuelas J, Pinol J, Ogaya R. Estimation of plant water concentration by the reflectance water index WI(R900/R970). *International Journal of Remote Sensing*, 1997, 18(13): 2869~2875.
- [ 11 ] A R Huete, H Q Liu, K Batchily. A comparison of vegetation indices over a global set of TM images for EOS-MODIS. *Remote Sensing of Environment*, 1997, 59: 440~451.
- [ 12 ] 金秀良, 徐新刚, 李振海, 等. 基于新型植被指数对冬小麦蛋白质含量的估算研究. *光谱学与光谱分析*, 2013, 33(9): 2541~2545.
- [ 13 ] Dash J, Curran P J. Evaluation of the MERIS terrestrial chlorophyll index(MTCI). *Advances in Space Research*, 2007, 39(1): 100~104.
- [ 14 ] Chen P F, Haboudane D, Tremblay N, et al. New spectral indicator assessing the efficiency of crop nitrogen treatment in corn and wheat. *Remote sensing of Environment*, 2010, 114: 1987.
- [ 15 ] Giteison. Wide dynamic range vegetation index for remote quantification of biophysical characteristics of vegetation. *Journal of Plant Physiology*, 2004, 161: 165~173.
- [ 16 ] Zhang Yanjiang. Interpretation of the modified soil-adjusted vegetation index isolines in red-NIR reflectance space. *Journal of Applied Remote Sensing*, 2007, 1: 1~12.
- [ 17 ] G Metternicht. Vegetation indices derived from high-resolution airborne videography for precision crop management. *International Journal of Remote Sensing*, 2003, 24: 2855.
- [ 18 ] Penuelas J. Assessment of photosynthetic radiation-use efficiency with spectral reflectance. *New Phytologist*, 1995, 131: 291.
- [ 19 ] Broge N H, Leblanc E. Comparing prediction power and stability of broad band and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sensing of Environment*, 2001, 76(2): 156~172.
- [ 20 ] Huete A R. A soil adjusted vegetation index SAVI. *Remote Sensing of Environment*, 1988, 25(3): 295~309.
- [ 21 ] Rhewl C, Vander S A, Kearney A, et al. Validation of the normalized difference vegetation index as a measure of neighborhood greenness. *Annals of Epidemiology*, 2011, 21(12): 946~952.
- [ 22 ] Clevers J G. Application of a weighted infrared-red vegetation index for estimating leaf area index by correcting for soil moisture. *Remote Sensing of Environment*, 1989, 29(1): 25~37.
- [ 23 ] Fitzgerald G, Rodriguez D. Measuring and predicting canopy nitrogen nutrition in wheat using a spectral index-the canopy chlorophyll content index(CCCI). *Field Crops Research*, 2010, 116(3): 318~324.
- [ 24 ] 束美艳, 顾晓鹤. 基于新型植被指数的冬小麦 LAI 高光谱反演. *中国农业科学*, 2018, 51(18): 3486~3496.
- [ 25 ] Penuelas J. Reflectance indices associated with physiological changes in nitrogen and water-limited sunflower leaves. *Remote Sensing Environment*, 1994, 48(2): 135~146.
- [ 26 ] Gamon J A, Penuelas J. A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sensing of Environment*, 1992, 41(1): 35~44.
- [ 27 ] Chen J M. Evaluation of vegetation indices and modified simple ratio for boreal application. *Canadian Journal of Remote Sensing*, 1996, 22(3): 229~242.
- [ 28 ] Fei Li, Martin L. Estimating N status of winter wheat using a handheld spectrometer in the North China Plain. *Field Crops Research*, 2008(106): 77~85.
- [ 29 ] Wang Fumin. New vegetation index and its application in estimating leaf area index of rice. *Rice Science*, 2007, 14(3): 195~203.
- [ 30 ] Jean-Louis Roujean, Francois-Marie Breon. Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. *Remote Sensing of Environment*, 1995, 51(3): 375~384.
- [ 31 ] 傅银贞. 遥感植被指数分析及应用研究. 福州: 福州大学, 2010.
- [ 32 ] Haboudane D, Miller J R, et al. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: modeling and validation in the context of precision agriculture. *Remote Sensing of Environment*, 2004, 90(3): 337~352.
- [ 33 ] Zhongsheng Cao, Xia Yao, et al. Comparison of the abilities of vegetation indices and photosynthetic parameters to detect heat stress in wheat. *Agricultural and Forest Meteorology*, 2019, 25: 121~136.
- [ 34 ] 管晓梅. 优化因子的偏最小二乘法及其在高光谱检测果糖含量中的应用. 济南: 山东师范大学, 2018.
- [ 35 ] 张海云, 刘桂兰. 组合预测方法及其在粮食产量预测中的应用. *农机化研究*, 2002, 11(4): 166~173.

## Estimation of winter harvest index based on PLS and combined prediction method

Chen Guo<sup>1, 2, 3</sup>, Xu Xingang<sup>2, 3\*</sup>, Du Xiaochu<sup>1</sup>, Yang Guijun<sup>2, 3</sup>, Zhao Xiaoqing<sup>2, 3</sup>,  
Wei Pengfei<sup>1, 2, 3</sup>, Wang Yulong<sup>2, 3</sup>, Fan Lingling<sup>2, 3</sup>

(1.College of Resources and Environment, Hubei University, Wuhan 430062, China; 2.Beijing Agricultural Information Technology Research Center, Beijing 100097, China; 3.National Agricultural Information Engineering Technology Research Center, Beijing 100097, China )

**Abstract:** [ **Purpose** ] Harvest index can effectively reflect the ability of crop population photohyalates to be transformed into grain accumulation. Also it is a key index to evaluate the yield level of crop varieties. Harvest index can be measured in practice. Remote sensing inversion can save time and manpower, but the accuracy needs to be improved. [ **Method** ] The canopy spectral data of winter wheat in several key growth stages were used to analyze the correlation between the selected 44 common vegetation indexes and the measured harvest indexes. Five optimal typical vegetation indexes are selected in each growth stage. Then, the partial least squares ( PLS ) modeling was applied to obtain the HI remote sensing estimation models based on the spectral information of single growth period. Finally, the combination prediction theory was used to apply the combination prediction method to assign the optimal weight to each HI spectral model for each single growth periods, and finally a HI spectral combination prediction model based on the data of multiple growth periods was constructed. [ **Result** ] ( 1 ) After using the PLS, the modeling results of single growth period have improved, compared with that of single vegetation index. However it still needs to be improved. ( 2 ) The HI combination prediction model based on the combination prediction principle significantly improved the estimation accuracy of HI, with R<sup>2</sup> up to 0.55, which was 13% higher than the modeling prediction of single growth period. [ **Conclusion** ] Based on the combination prediction method of information of multiple growth stages, the optimal weight was given to the HI prediction model of each single growth stage to optimize the combination. In essence, the contribution of each growth stage to the HI formation of crops was indirectly utilized to significantly improve the estimation accuracy of winter wheat harvest index. It is a novel remote sensing estimation method of crop HI.

**Key words:** harvest index; PLS; winter wheat; combined forecasting model; remote sensing spectrum; growth durations