

## ИНФОРМАТИКА

УДК 539.3

*В. М. Буре*<sup>1,2</sup>, *Е. В. Канаши*<sup>1</sup>, *О. А. Митрофанова*<sup>1,2</sup>**АНАЛИЗ ХАРАКТЕРИСТИК ЦВЕТА РАСТЕНИЙ  
ПО АЭРОФОТОСНИМКАМ С РАЗЛИЧНЫМИ ФАКТОРАМИ  
КАЧЕСТВЕННЫХ ПОКАЗАТЕЛЕЙ**

<sup>1</sup> Агрофизический научно-исследовательский институт, Российская Федерация, 195220, Санкт-Петербург, Гражданский пр., 14

<sup>2</sup> Санкт-Петербургский государственный университет, Российская Федерация, 190034, Санкт-Петербург, Университетская наб., 7–9

Одним из актуальных и перспективных направлений в современном точном земледелии является оценка состояния растений. Точная оценка состояния сельскохозяйственных растений в период вегетации необходима для эффективного использования удобрений, выгодной урожайности и высокого качества продукции. Метод решения этой задачи основан на анализе цветовых характеристик растений по цифровым изображениям. В работе рассмотрены методы анализа характеристик цвета растений по аэрофотоснимкам с разными факторами качественных показателей. Кроме того, представлен пример анализа экспериментальных данных с помощью языка программирования R. Исходными данными задачи являются цветовые параметры  $L$ ,  $a$ ,  $b$  растений на специальных тестовых площадках. Тестовая площадка представляет собой небольшой участок поля, на котором известны качественные показатели растений. Описан следующий пример: имеются тестовые площадки пшеницы с известными дозами азота (0, 60, 90, 120 кг действующего вещества на 1 га). Кроме того, на каждой площадке сформированы определенные качественные показатели растений: размер зерна (крупные, мелкие), защита растений (сорняки, без сорняков), нормы высева (6 млн на 1 га, 5 млн на 1 га). Необходимо проанализировать существование линейной зависимости между цветом растений и дозой азота в зависимости от разных качественных факторов. В ходе решения задачи разработаны и протестированы алгоритмы для реализации представленных методов. В результате предварительного анализа в описанном примере распределения выборок характеристик цвета по каждой паре факторов оказались различными. В ходе эксперимента были построены 8 линейных регрессий, преимущественно уравнения регрессий оказались статистически значимыми в целом. Тем не менее следует отметить, что коэффициент  $\alpha$  при составляющей цвета  $L$  оказался равен 0. Предположительно это связано с ошибками во время закладки эксперимента (некоторая часть тестовых площадок была заложена позже остальных). Библиогр. 5 назв. Табл. 1.

*Ключевые слова:* аэрофотосъемка, обобщенная характеристика цвета, точное земледелие, язык R.

*Буре Владимир Мансурович* — доктор технических наук, профессор; vlb310154@gmail.com

*Канаши Елена Всеволодовна* — доктор биологических наук; ykanash@yandex.ru

*Митрофанова Ольга Александровна* — аспирант; omitrofa@gmail.com

*Bure Vladimir Mansurovich* — doctor of technical sciences, professor; vlb310154@gmail.com

*Kanash Elena Vsevolodovna* — doctor of biological sciences; ykanash@yandex.ru

*Mitrofanova Olga Aleksandrovna* — postgraduate student; omitrofa@gmail.com

© Санкт-Петербургский государственный университет, 2017

## ANALYSIS OF PLANTS COLOR CHARACTERISTICS USING AEROPHOTOS WITH DIFFERENT FACTORS OF QUALITATIVE INDICATORS

<sup>1</sup> Agrophysical research institute, 14, Grazhdanskiy pr., St. Petersburg, 195220, Russian Federation

<sup>2</sup> St. Petersburg State University, 7–9, Universitetskaya nab., St. Petersburg, 199034, Russian Federation

One of the most relevant and highly demanded directions in modern precision agriculture is the assessment of vegetation conditions. An accurate assessment of the state of agricultural plants during the growing season is necessary for the effective use of fertilizers, profitable yields and high quality products. The method for solving this problem is based on an analysis of the color characteristics of plants from digital images. In this paper methods of the analysis of color characteristics of plants in aerial photos with various factors of qualitative indicators are examined. In addition, an example of analysis of experimental data is presented using the programming language R. The initial data of the problem are plants' color parameters  $L$ ,  $a$ ,  $b$  in special test areas. The test area is a small region of the field where the qualitative indices of plants are already known. In this paper, the following example is considered: there are test areas of wheat with known doses of nitrogen (0, 60, 90, 120 kg of active substance per 1 ha). In addition, certain quality indicators of plants are formed at each site: grain size (large, small), plant protection (weeds, no weeds), seeding rates (6 mmillion per ha, 5 mmillion per ha). The existence of a linear relationship between the color of plants and the dose of nitrogen must be analysed on the basis of various qualitative factors. In the course of solving the problem, algorithms were developed and tested to implement the methods presented. As a result of a preliminary analysis in the example described, the distribution of samples of color characteristics for each pair of factors turned out to be different. In the course of the experiment 8 linear regressions were developed and the regression equations turned out to be statistically significant as a whole. Nevertheless, it should be noted that the coefficient  $\alpha$  for the color component of  $L$  turned out to be 0. Presumably, this is due to errors during the experiment's stowage (some of the test areas were laid out later than others). Refs 5. Table 1.

*Keywords:* aerial photography, generalized color characteristic, precision agriculture, language R.

**Введение.** В настоящее время при исследовании задач, связанных с точным земледелием и экологией, все более актуальным и перспективным направлением становится применение различных методов статистического анализа данных, а также методов анализа цифровых изображений. Эти методы используются для решения следующих наиболее востребованных задач:

- обнаружение чрезвычайных ситуаций;
- мониторинг состояния растений;
- прогноз урожайности;
- дифференцированное внесение удобрений;
- мониторинг состояния систем мелиорации и др.

Для решения перечисленных проблем могут использоваться разные статистические методы: бинарная регрессия [1, 2], кригинг [3], построение калибровочных кривых [4] и т. д.

В данной работе представлены методы статистического анализа характеристик цвета растений с различными факторами стресса по аэрофотоснимкам. Этот анализ необходим для исследования возможностей последующей оценки состояния растений на любом участке поля по данным аэрофотосъемки. Для интерпретации цветовых параметров применяется цветовая модель CIELAB.

**Объекты и методы.** Для получения аэрофотоснимков использовался автоматизированный беспилотный летательный аппарат (БЛА) — квадрокоптер Геоскан-401 [5]. Полезной нагрузкой служили две зеркальные камеры, позволяющие дать высококачественные изображения без каких-либо цифровых шумов в видимом и ближнем инфракрасном спектрах.

Весь процесс получения аэрофотоснимков с помощью автоматизированного БЛА можно представить следующими основными этапами.

1. План полета. В поставленную комплектацию БЛА Геоскан-401 также входит специальное программное обеспечение GeoScan Planner, позволяющее определять границы полей, которые необходимо исследовать (например, с помощью карт Google Earth), после чего план полета создается автоматически. В результате созданный план загружается в БЛА.

2. Полет и съемка. Данный этап полностью автоматизирован, необходимо только запустить БЛА на заданный маршрут, и система будет автоматически делать снимки с помощью бортовых камер и датчиков во время полета. Для съемки квадрокоптер использует координаты GPS, при этом основное требование для съемки — достаточное перекрытие изображений.

3. Предварительная обработка изображений. В комплект поставки также входит программное обеспечение Agisoft PhotoScan, которое дает возможность автоматически создавать ортофотопланы в формате Geotiff из мозаики полученных изображений, а также трехмерные рельефные модели.

Объектами исследования являются опытные сельскохозяйственные поля, расположенные на территории филиала Агрофизического института в дер. Меньково Ленинградской обл.

Исходными данными задачи являются цветовые параметры  $L$ ,  $a$ ,  $b$  растений на тестовых площадках (параметр  $L$  всегда положителен и обозначает светлоту;  $a > 0$  обозначает параметр красной составляющей,  $a < 0$  — зеленой,  $b > 0$  — желтой,  $b < 0$  — синей). Тестовая площадка представляет собой небольшой участок поля, на котором известны качественные показатели растений. Проведем обследование следующего примера: имеются тестовые площадки пшеницы с известными дозами азота (0, 60, 90, 120 кг действующего вещества на 1 га). Кроме того, на каждой площадке сформированы определенные качественные показатели растений: размер зерна (крупные, мелкие), защита растений (сорняки и без них), нормы высева (6 млн на 1 га, 5 млн на 1 га). Необходимо проанализировать существование линейной зависимости между цветом растений и дозой азота в зависимости от различных качественных факторов.

В первую очередь необходимо представить характеристики цвета в виде единой величины. В связи с этим предложено свернуть вектор параметров цвета в одно значение с использованием выпуклой линейной комбинации. Данное значение условно названо обобщенной характеристикой цвета (generalized color characteristic)  $C_{Lab}$ , которая рассчитывается по формуле [4]

$$C_{Lab} = \alpha L + \beta_1 a + \beta_2 b, \quad (1)$$

где  $\alpha \geq 0$ ,  $\beta_1 \geq 0$  и  $\beta_2 \geq 0$  — эмпирически подбираемые коэффициенты для каждого цифрового изображения посева, причем

$$\alpha + \beta_1 + \beta_2 = 1.$$

Основной критерий подбора коэффициентов  $\alpha$ ,  $\beta_1$ ,  $\beta_2$  — обеспечение максимально выраженной линейной связи между характеристикой цвета (1) и дозой азота на тестовых площадках. Для этого в ходе решения выполняется построение восьми уравнений линейных регрессий (для каждого набора факторов). При таком подходе должно быть выполнено следующее условие для обеспечения наиболее выраженной линейной связи:

$$\max_{(\alpha, \beta_1)} \min_{i=1, \dots, 8} R_i^2(\alpha, \beta_1), \quad (2)$$

где  $R_i^2$  — коэффициенты детерминации построенных линейных регрессий.

Наиболее естественным путем решения этой задачи представляется перебор коэффициентов  $\alpha$ ,  $\beta_1$ ,  $\beta_2$ . Соответственно для каждого набора коэффициентов выполняются построение уравнений линейных регрессий, вычисление коэффициентов детерминации и поиск максимума (2).

**Результаты и их обсуждение.** Предварительно целесообразно провести анализ распределений выборок характеристик цвета по каждому фактору. В случае, если соответствующая пара факторов одновременно будет иметь нормальное распределение, то можно провести их сравнение параметрическим методом, а если она одновременно будет иметь ненормальное распределение, то можно применить непараметрический метод. Алгоритм основной функции параметрического метода схематично представим следующим образом:

// Input data:

**Data:**  $L^i = (L_1^i, \dots, L_n^i)$ ,  $a^i = (a_1^i, \dots, a_n^i)$ ,  $b^i = (b_1^i, \dots, b_n^i)$ ,  $i = 1, 2$ , — цветовые параметры растений на тестовых площадках для каждого парного фактора;  $n$  — количество тестовых площадок.

// Output data:

**Result:** Эмпирически подобранные коэффициенты  $\alpha$ ,  $\beta_1$ ,  $\beta_2$ ; сообщение о результатах сравнения распределений.

//

min  $\leftarrow 1$ ;

bol  $\leftarrow$  FALSE;

**for**  $\alpha \in [0; 1]$  **do**

**for**  $\beta_1 \in [\alpha; 1]$  **do**

    {

$\beta_2 \leftarrow 1 - \alpha - \beta_1$ ;

        Вычисление обобщенной характеристики цвета для первого парного фактора

$C_{Lab}^1 = (C_{Lab,1}^1, \dots, C_{Lab,n}^1)$ ;

        Вычисление обобщенной характеристики цвета для второго парного фактора

$C_{Lab}^2 = (C_{Lab,1}^2, \dots, C_{Lab,n}^2)$ ;

        Критерий Фишера с помощью R-функции **var.test**;

**if** p.value  $< 0.05$  **then**

**if** p.value  $<$  min **then**

            {

                min  $\leftarrow$  p.value;

                Запоминание коэффициентов  $\alpha$ ,  $\beta_1$ ,  $\beta_2$ ;

            }

**else if** bol == FALSE **then**

```

        if Статистика входит в доверительный интервал then
            bol ← TRUE;
    }
if min < 1 then
    Return min,  $\alpha$ ,  $\beta_1$ ,  $\beta_2$ , сообщение "дисперсии не равны";
if min == 1 then
    if bol == TRUE then
        Критерий Стьюдента (возвращает результат сравнения средних);
    else сообщение об ошибке.

```

Алгоритм основной функции непараметрического метода можно представить следующим образом:

```

// Input data:
Data:  $L^i = (L_1^i, \dots, L_n^i)$ ,  $a^i = (a_1^i, \dots, a_n^i)$ ,  $b^i = (b_1^i, \dots, b_n^i)$ ,  $i = 1, 2$ , — цветовые параметры растений на тестовых площадках для каждого парного фактора;  $n$  — количество тестовых площадок.
// Output data:
Result: Эмпирически подобранные коэффициенты  $\alpha$ ,  $\beta_1$ ,  $\beta_2$ ; сообщение о результатах сравнения распределений.
//
min ← 1;
for  $\alpha \in [0; 1]$  do
    for  $\beta_1 \in [\alpha; 1]$  do
        {
             $\beta_2 \leftarrow 1 - \alpha - \beta_1$ ;
            Вычисление обобщенной характеристики цвета для первого парного фактора
             $C_{Lab}^1 = (C_{Lab,1}^1, \dots, C_{Lab,n}^1)$ ;
            Вычисление обобщенной характеристики цвета для второго парного фактора
             $C_{Lab}^2 = (C_{Lab,1}^2, \dots, C_{Lab,n}^2)$ ;
            Критерий Вилкоксона с помощью R-функции wilcox.test;
            if p.value < 0.05 then
                if p.value < min then
                    {
                        min ← p.value;
                        Запоминание коэффициентов  $\alpha$ ,  $\beta_1$ ,  $\beta_2$ ;
                    }
                }
        }
if min < 1 then
    Return min,  $\alpha$ ,  $\beta_1$ ,  $\beta_2$ , сообщение "распределения различны";
if min == 1 then
    Сообщение "распределения одинаковы".

```

В результате проведенного анализа распределения всех пар факторов оказались различными. Это позволяет предполагать, что в дальнейшем существует возможность классификации растений по определенному фактору.

Для решения основной задачи был разработан алгоритм, выполняющий построение линейных регрессионных моделей, а также проверку их значимости. Схематично

его можно представить следующим образом:

// Input data:

**Data:**  $N = (N_1, \dots, N_n)$  — величины доз азота на тестовых площадках;  $n$  — количество тестовых площадок;  $L^i = (L_1^i, \dots, L_n^i)$ ,  $a^i = (a_1^i, \dots, a_n^i)$ ,  $b^i = (b_1^i, \dots, b_n^i)$ ,  $i = 1, \dots, 8$ , — соответствующие им цветовые параметры для каждого набора факторов.

// Output data:

**Result:** Эмпирически подобранные коэффициенты  $\alpha$ ,  $\beta_1$ ,  $\beta_2$ ; уравнения построенных линейных регрессий; коэффициенты детерминации  $R_i^2$ ,  $i = 1, \dots, 8$ ; стандартные ошибки;  $t$ -значения;  $F$ -статистики.

//

max  $\leftarrow 0$ ;

**for**  $\alpha \in [0; 1]$  **do**

**for**  $\beta_1 \in [\alpha; 1]$  **do**

    {

$\beta_2 \leftarrow 1 - \alpha - \beta_1$ ;

        Вычисление обобщенных характеристик цвета

$C_{Lab}^i = (C_{Lab,1}^i, \dots, C_{Lab,n}^i)$ ,  $i = 1, \dots, 8$ ;

**if** Значения  $C_{Lab}^i$  монотонны **then**

        {

            Построение первой линейной регрессионной модели

$C_{Lab}^1 = b_0^1 + b_1^1 N + \epsilon^1$  с помощью R-функции **lm**;

            Вычисление коэффициента детерминации  $R_1^2$ ;

            min  $\leftarrow R_1^2$ ;

            Построение остальных линейных регрессионных моделей

$C_{Lab}^i = b_0^i + b_1^i N + \epsilon^i$ ,  $i = 2, \dots, 8$ ;

            Вычисление коэффициентов детерминации  $R_i^2$ ,  $i = 2, \dots, 8$ ;

**if**  $R_i^2 < \min$  **then**

                min  $\leftarrow R_i^2$ ;

**if** min  $>$  max **then**

            {

                max  $\leftarrow \min$ ;

                Запоминание коэффициентов  $\alpha, \beta_1, \beta_2$ ;

            }

        }

    }

Return: max,  $\alpha, \beta_1, \beta_2$ ;

Проверка адекватности построенных линейных регрессионных моделей.

В результате анализа экспериментальных данных были построены 8 линейных регрессий, преимущественно уравнения регрессий оказались статистически значимыми в целом. Тем не менее следует отметить, что коэффициент  $\alpha$  при составляющей цвета  $L$  оказался равен 0. Предположительно это связано с ошибками во время закладки эксперимента (некоторая часть тестовых площадок была заложена позже остальных). Выходные данные проведенного эксперимента приняли следующие значения:  $\max_{(\alpha, \beta_1)} \min_{i=1, \dots, 8} R_i^2(\alpha, \beta_1) = 0.4312$ ,  $\alpha = 0$ ,  $\beta_1 = 0.61$ ,  $\beta_2 = 0.39$ . Результаты построения линейных регрессий представлены в таблице.

## Результаты построения линейных регрессий

№ анализа	Набор факторов (норма высева — млн на 1 га, размер зерна, наличие сорняков)	Степень значимости коэффициента 1	Степень значимости коэффициента 2	Коэффициент детерминации	p-value
1	5, крупные, сорняки	0.861	0.000008	0.8759	0.000008
2	5, крупные, чистые	0.01973	0.00596	0.5472	0.005957
3	5, мелкие, сорняки	0.238086	0.000646	0.7039	0.000646
4	5, мелкие, чистые	0.000013	0.0204	0.4312	0.02037
5	6, крупные, сорняки	0.0000001	0.00000005	0.9549	0.00000005
6	6, крупные, чистые	0.0000001	0.00000003	0.9588	0.00000003
7	6, мелкие, сорняки	0.00187	0.00316	0.5983	0.003163
8	6, мелкие, чистые	0.0000027	0.0167	0.4515	0.0167

**Заключение.** В работе представлены методы анализа цветковых характеристик растений по аэрофотоснимкам, разработаны и протестированы алгоритмы для проведения анализа. На основе проведенного исследования можно сформулировать следующие выводы:

- 1) предложенные методы анализа позволят в дальнейшем осуществить оценку состояния растений на любом участке поля, что необходимо прежде всего для эффективного использования удобрений, повышения урожайности и качества продукции, а также для улучшения экологического состояния сельскохозяйственного поля;
- 2) целесообразно повторить эксперимент, увеличив количество тестовых площадок, а также исключив ошибки при их закладке;
- 3) следует провести ковариационный анализ построенных линейных регрессионных моделей, а также изучить возможность классификации растений по факторам качественных показателей;
- 4) необходимо определить возможность калибровки данных на основе тестовых площадок для последующей оценки состояния растений по факторам качественных показателей в регионе (Ленинградская обл.).

### Литература

1. Буре В. М. Методология применения бинарной регрессии в точном земледелии // Математические модели в теоретической экологии и земледелии: материалы Междунар. семинара, посвященного памяти профессора Ратмира Александровича Полуэктова (Полуэктовские чтения). 2014. С. 118–121.
2. Якушев В. П., Буре В. М., Париллина Е. М. Бинарная регрессия и ее применение в агрофизике. СПб.: Агрофиз. ин-т, 2015. 36 с.
3. Буре В. М., Митрофанова О. А. Прогноз пространственного распределения экологических данных с применением кригинга и бинарной регрессии // Вестн. С.-Петербург. ун-та. Сер. 10. Прикладная математика. Информатика. Процессы управления. 2016. Вып. 3. С. 97–105.
4. Митрофанова О. А., Буре В. М., Канааш Е. В. Математический модуль для автоматизации колориметрического метода оценки обеспеченности растений азотом // Вестн. С.-Петербург. ун-та. Сер. 10. Прикладная математика. Информатика. Процессы управления. 2016. Вып. 1. С. 85–91.
5. БЛА Геоскан-401 // URL: <https://www.geoscan.aero> (дата обращения: 01.04.2017).

**Для цитирования:** Буре В. М., Канааш Е. В., Митрофанова О. А. Анализ характеристик цвета растений по аэрофотоснимкам с различными факторами качественных показателей // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. 2017. Т. 13. Вып. 3. С. 278–285. DOI: 10.21638/11701/spbu10.2017.305

### References

1. Bure V. M. Metodologija primeneniiia binarnoi regressii v tochnom zemledelii [Methodology of using binary regression in precision agriculture]. *Mathematical models in theoretical ecology and*

*agriculture: materials of Intern. Seminar dedicated to the memory of Professor Ratmir Alexandrovich Poluektov (Poluektov's reading)*, 2014, pp. 118–121. (In Russian)

2. Yakushev V. P., Bure V. M., Parilina E. M. *Binarnaia regressiia i ee primenenie v agrofizike [Binary regression and its application in agrophysics]*. Saint Petersburg, Agrophys. Institute, 2015, 36 p. (In Russian)

3. Bure V. M., Mitrofanova O. A. Prognoz prostranstvennogo raspredeleniia ekologicheskikh dannykh s primeneniem kriginga i binarnoi regressii [Prediction of the spatial distribution of ecological data using kriging and binary regression]. *Vestnik of Saint Peterburg University. Series 10. Applied Mathematics. Computer Science. Control Processes*, 2016, iss. 3, pp. 97–105. (In Russian)

4. Mitrofanova O. A., Bure V. M., Kanash E. V. Matematicheskii modul' dlia avtomatizatsii kolorimetriceskogo metoda otsenki obespechennosti rastenii azotom [Mathematical module to automate the colorimetric method for estimating nitrogen status of plants]. *Vestnik of Saint Peterburg University. Series 10. Applied mathematics. Computer Science. Control Processes*, 2016, iss. 1, pp. 85–91. (In Russian)

5. *UAV Geoscan-401*. Available at: <https://www.geoscan.aero> (accessed: 01.04.2017).

**For citation:** Bure V. M., Kanash E. V., Mitrofanova O. A. Analysis of plants color characteristics using aerophotos with different factors of qualitative indicators. *Vestnik of Saint Petersburg University. Applied Mathematics. Computer Science. Control Processes*, 2017, vol. 13, iss. 3, pp. 278–285. DOI: 10.21638/11701/spbu10.2017.305

Статья рекомендована к печати проф. Л. А. Петросяном.

Статья поступила в редакцию 27 апреля 2017 г.

Статья принята к печати 8 июня 2017 г.