Development of a mathematical model for forecasting pastural fertility to remote sensing data

V Trukhachev, S Oliinyk, T Lesnyak, D Litvin and N Zlydnev
Stavropol State Agrarian University, Russia
E-mail: soliynik60@gmail.com

Abstract. Variants of developing a mathematical model for predicting pasture fertility based on the use of remote sensing data from the Earth and ground based measurements on test field plots are presented. The NDVI data, which is a time series with pronounced seasonal features in the form of one or two maxima, are the basis for obtaining the vegetation index forecast model (NDVI) from the currently available measurements. To predict the parameters of this maximum, we used the position on the time axis, width, amplitude, and its shape. Using various model regressions over the entire available time interval, it is possible to obtain models, for example, in the form of rational and polynomial functions. As model, in addition to the rational and polynomial ones presented above, we used the following functions with characteristic impulse behavior: Cauchy, arctg (x) ', Verhulst, Gompertz differential functions (derivatives) with left and right asymmetries, as well as the sum fs and the product fp of logistic functions Verhulst. Of all the options on the evaluation interval [56 and 180] days, the model in the form of the sum of the Verhulst logistic functions with 7 evaluated parameters turned out to be more effective. The results show that all parameters are statistically significant, and the corrected determination coefficient is large enough: 0.9545. However, to parameterize the model, observations are needed “on the decline” of the time series - 150-180 days from the beginning of the year. The use of mathematical models is necessary for managing pasture plots and the need for timely management decisions of various kinds - alternating corrals, mowing terms for grass stands, and fertilizing. To generate reliable information about the state of pastures, the use of remote monitoring is relevant.

1. Introduction
A modern, innovative approach to the development of animal husbandry should provide high predictability of the final result, information support in informed management decisions, reducing the average possible risks, and ultimately, increasing labor productivity in this important area for the country [1-2].

The availability of modern information-measuring systems for aerospace pasture monitoring, as well as relevant information technologies, require further improvement of predictive mathematical models of both pasture productivity and productive qualities of pasture animals [3].

Other things being equal, the indicated characteristics in the operational time scale are determined mainly by weather and climatic conditions.

The aim of our research was to develop a mathematical model for predicting pasture fertility based on remote sensing data [4-5].
2. Materials and methods
In our work we used the deviation of the observed factor characteristics from their changes averaged over the years of observation to build a model for predicting pasture fertility. Such deviations in a particular year obviously contain climatic features of a given year, which should be used for forecasting.

To determine an adequate model of temperature changes during the year, model regressions were considered in the form of trigonometric series of the 1st, 2nd and 3rd orders, as well as models of the moving average (smooth) with different periods of averaging.

Creating a model for predicting the biological productivity of the pasture and the nutritional value of the feed received from it from the temperature trajectory of the season, the amount of precipitation, as well as from the levels of the normalized vegetation index NDVI (Normalized difference vegetation index) recorded using satellites [6-7].

3. Results and Discussion
To solve the problem of developing a mathematical model for predicting pasture fertility, the vegetative index NDVI was used (figure 1).

![Figure 1](image1.png)

Figure 1. Here along the axis OX - time in days. Countdown – 01.01.2018.

At the first stage, we tried to obtain a forecast model for the NDVI index based on the currently available measurements. NDVI data represent a time series with pronounced seasonal features in the form of one or two maxima. It is necessary to predict the parameters of this maximum - position on the time axis, width, amplitude and its shape.

Using various model regressions on the existing time interval, it is possible to obtain models, for example, in the form of rational and polynomial functions (figure 2).

![Figure 2](image2.png)

Figure 2. Rational and polynomial function.

On the left side are the equations of the models, optimal on method least squares values of their parameters, and also in parentheses, 95% confidence intervals for the values of these parameters.

On the right side:
Sse – the sum of squared regression errors;
Rsquare – the coefficient of determination ($R^2$);
Dfe – the number of degrees of freedom of regression errors;
Adjrsquare – corrected coefficient of determination ($R_{adj}^2$);
Rmse – the standard error of the regression.

Linear model Poly6:
$$fpoly6(x) = p1*x^6 + p2*x^5 + p3*x^4 + p4*x^3 + p5*x^2 + p6*x + p7$$
where $x$ is normalized by mean 193.8 and std 68.92
Coefficients (with 95% confidence bounds):
p1 = 0.05506 (0.02669, 0.08342)
p2 = 0.07389 (0.03773, 0.1101)
p3 = -0.2473 (-0.3757, -0.1189)
p4 = -0.109 (-0.231, 0.01298)
p5 = 0.353 (0.196, 0.5101)
p6 = -0.1487 (-0.241, -0.05632)
p7 = 0.2201 (0.1758, 0.2644)

Sse: 0.0933
Rsquare: 0.8588
Dfe: 25
Adjrsquare: 0.8250
Rmse: 0.0611

Linear model Poly3:
$$fpoly3(x) = p1*x^3 + p2*x^2 + p3*x + p4$$
where $x$ is normalized by mean 193.8 and std 68.92
Coefficients (with 95% confidence bounds):
p1 = 0.1209 (0.08919, 0.1527)
p2 = 0.05048 (0.01627, 0.0847)
p3 = -0.275 (-0.3447, -0.2054)
p4 = 0.277 (0.2323, 0.3216)

Sse: 0.2134
Rsquare: 0.6772
Dfe: 28
Adjrsquare: 0.6426
Rmse: 0.0873

frat22(x) = ($p1*x^2 + p2*x + p3$) / ($x^2 + q1*x + q2$)
Coefficients (with 95% confidence bounds):
p1 = 0.0819 (0.04805, 0.1158)
p2 = 4.021 (-0.4181, 8.461)
p3 = -500 (fixed at bound)
q1 = -164.7 (-185.2, -144.3)
q2 = 7819 (5723, 9916)

Sse: 0.1176
Rsquare: 0.8200
Dfe: 28
Adjrsquare: 0.8030
Rmse: 0.0648

At the same time, it can be seen that of the above models, only the third-order polynomial model is statistically significant (with a reliability of 95%). A similar result is obtained for other so-called custom models.

If we limit the time interval on the right, simulating the limited nature of observations for forecasting problems, then it becomes completely impossible to obtain any meaningful model. This statement corresponds to the well-known requirement for the volume of statistical data - for each estimated parameter of the model should be at least 10 independent observations.

As one of the possible ways out of the situation, it is possible to propose a not quite methodically correct way. Let's try to generate "observations" dependent on existing ones, for example, using piecewise linear interpolation. In the following figure, such "observations" are indicated by blue stars (figure 3).
In addition to the above rational and polynomial functions, the following functions with characteristic impulse behavior were used as model ones: Cauchy, arctg (x)', Verhulst, Gompertz differential functions (derivatives) with left and right asymmetries, as well as the sum fs and the product fp of logistic functions Verhulst (figure 4).

Of all the options on the assessment interval [56 and 180] days or [02.27.2018, 06.27.2018], the “best” model was the sum of the Verhulst logistic functions with 7 evaluated parameters (figure 5).
Figure 5. Logistic Verhulst function with 7 evaluated parameters.

\[ fs(t) = C + \frac{k_d}{1 + \exp(m_d(t-t_0d))} + \frac{k_u}{1 + \exp(-mu*(t-t_0u))} \]

Coefficients (with 95% confidence bounds):
- \( C = (-0.3469, -0.2096) \)
- \( k_d = (0.2371, 0.2936) \)
- \( k_u = (0.481, 0.607) \)
- \( m_d = (0.1908, 0.4609) \)
- \( mu = (0.09932, 0.1523) \)
- \( t_0d = (155, 158) \)
- \( t_0u = (74.01, 78.77) \)

\[ sse: 0.1018 \]
\[ rsquare: 0.9573 \]
\[ dfe: 93 \]
\[ adjrsquare: 0.9545 \]
\[ rmse: 0.0331 \]

It can be seen that all parameters are statistically significant, and the corrected determination coefficient is quite large: 0.9545.

However, in this case, for parameterization of the model, observations are needed “on the decline” of the time series - 150-180 days from the beginning of the year or 02.06.2018-02.07.2018.

4. Conclusion

Based on the foregoing:

- The results show that all parameters are statistically significant, and the corrected determination coefficient is large enough: 0.9545. However, to parameterize the model, observations are needed “on the decline” of the time series - 150-180 days from the beginning of the year.
- The use of mathematical models is necessary for managing pasture plots and the need for timely management decisions of various kinds - alternating corrals, mowing terms for grass stands, and fertilizing. For the formation of reliable information about the state of pastures, the use of remote monitoring is relevant.

References

[1] Kashkin V B and Sukhinin A I 2008 Digital processing of aerospace images Retrieved from: http://files.lib.sfu-kras.ru/ebibl/umkd/54/u_course.pdf
[2] Zhelezova S V, Ananiev A A, Vyunov M V and Berezovsky E V 2016 Monitoring of winter wheat crops using unmanned aerial photography and an optical sensor GreenSeeker RT200. Bulletin of the Orenburg State University 6 56-61
[3] Obiralov A I, Limonov A N and Gavrilova L A 2006 Photogrammetry and remote sensing (Moscow: Kolos) 335
[4] Cherepanov A S and Druzhinina E G 2009 Spectral properties of vegetation and vegetation indices. Geomatics 3 28-32
[5] Lugassi R, Zaady E, Chudnovsky A, Dvash L and Goldshleger N 2015 Spectral Slope as an
Indicator of Pasture Quality. *Remote Sens* **7**(1) 256–74

[6] Landsat Data Continuity Mission Retrieved from: www.nasa.gov/pdf/723395main_LDCM_presskit2013-final.pdf

[7] Antonov V N and Sladkikh L A 2009 Monitoring of the state of crops and forecasting the yield of spring wheat according to remote sensing data. Monitoring the state of crops and forecasting the yield of spring wheat according to the data DZZ. *Geomatics* **4** 50–3

[8] Nosov V V, Kozin M N, Andreev V I, Surzhanskaya I Y and Murzina E A 2016 Increasing The Efficiency of Land Resources Use for an Agricultural Enterprise Research. *Journal of Pharmaceutical, Biological and Chemical Sciences* **7**(6) 382–385

[9] Zhichkin K, Nosov V, Zhichkina L, Tkachev S and Voloshchuk L 2020 Prediction methodology for potential damage from misuse of agricultural lands. *E3S Web of Conferences* **161** 01060

[10] Zhichkin K, Nosov V, Zhichkina L, Zhenzhebir V and Rubtsova S 2020 The agricultural crops production profitability in modern conditions. *E3S Web of Conferences* **175** 13008