HAL
open science

Descriptive modeling to predict deoxynivalenol in winter wheat in The Netherlands

Hj van Der Fels-Klerx, Saskia Lge Burgers, Cjh Booij

To cite this version:

Hj van Der Fels-Klerx, Saskia Lge Burgers, Cjh Booij. Descriptive modeling to predict deoxynivalenol in winter wheat in The Netherlands. Food Additives and Contaminants, 2010, 27 (05), pp.636-643. 10.1080/19440040903571762 . hal-00591168

HAL Id: hal-00591168

https://hal.science/hal-00591168v1

Submitted on 7 May 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Descriptive modeling to predict deoxynivalenol in winter wheat in The Netherlands

Journal:	*Food Additives and Contaminants*
Manuscript ID:	TFAC-2009-318.R1
Manuscript Type:	Special Issue
Methods/Techniques:	Risk assessment - modelling, Regression, Statistical analysis
Additives/Contaminants:	Mycotoxins, Mycotoxins - trichothecenes, Mycotoxins - fusarium
Food Types:	Cereals and grain

SCHOLARONE™
Manuscripts

For Peer Review Only

Running head: Predictive modeling of deoxynivalenol in wheat

Descriptive modeling to predict deoxynivalenol in winter wheat in The Netherlands

H J VAN DER FELS-KLERX[a,*], SLGE BURGERS[a,b] and CJH BOOIJ[b]

[a]*RIKILT - Institute of Food Safety, Wageningen UR, PO Box 230, NL-6700 AE Wageningen, The Netherlands*; [b]*Plant Research International, Wageningen UR, PO Box 16, NL-6700 AA Wageningen, The Netherlands*

Abstract

Predictions of deoxynivalenol (DON) content in wheat at harvest can be useful for decision making by stakeholders of the wheat feed and food supply chain. The objective of the current research was to develop quantitative predictive models for DON in mature winter wheat in The Netherlands for two specific groups of end-users. One model was developed for use by farmers in underpinning *Fusarium* spp. disease management, specifically the application of fungicides around wheat flowering (model A). The second model was developed for industry and food safety authorities, and considered the entire wheat cultivation period (model B). Model development was based on observational data collected from 425 fields throughout the Netherlands

*Corresponding author: HJ van der Fels-Klerx. E-mail: ine.vanderfels@wur.nl

For Peer Review Only

in the period 2001-2008. For each field, agronomical information, climatic data and DON levels in mature wheat were collected. Using multiple regression analyses, the set of biological relevant variables that provided the highest statistical performance was selected. The two final models include the following variables: region, wheat resistance level, spraying, flowering date, several climatic variables in the different stages of wheat growing, and length of the period between flowering and harvesting (model B only). The percentage of variance accounted for was 64.4 % and 65.6 % for model A and B, respectively. Model validation showed high correlation between the predicted and observed DON levels. The two models may be applied by various groups of end-users to reduce DON contamination in wheat derived feed and food products and, ultimately, reduce animal and consumer health risks.

Keywords: deoxynivalenol; mycotoxins; winter wheat; predictive modeling

For Peer Review Only

Introduction

Mycotoxins are secondary metabolites of fungi, some of which pose a threat to animal and human health. These toxins have been implicated as causative agents in a variety of animal disorders, such as depressed feed intake and feed refusal, decreased growth and feed efficiency, reproductive disorders, and skin and gastrointestinal irritation (Diekman and Green 1992; D'Mello et al. 1999; Morgavi and Riley 2007; Minervini and Aquila 2008). In human, these compounds have been associated with, amongst others, immunosurpression and cancer induction (Creppy 2002; Smith et al. 2006). Under favorable conditions, related to agronomical and environmental circumstances, fungi may colonize crops and produce mycotoxins. In wheat cultivation, the most prominent fungi are *Fusarium* spp., resulting in Fusarium Head Blight (FHB), one of the most important fungal diseases affecting all cropping areas of the world (Botallico and Perrone 2002; Champeil et al. 2004a; Osborne and Stein 2007). FHB results into reduced wheat yield and quality, and consequent economic losses. One of the most common toxins produced by *Fusarium* spp. is deoxynivalenol (DON) (Miller 2008). Besides the losses due to reduced wheat yield and quality, DON contaminated wheat has lower market value due to its potential health threats. Growing wheat is, therefore, faced with the challenge of keeping contamination with *Fusarium* spp. and related mycotoxins at the minimum. Application of Good Agricultural Practices (GAP) during cultivation, and Good Manufacturing Practices (GMP) and Hazard Analysis and Critical Control Point (HACCP) procedures during consecutive stages of the wheat feed and food supply chain will ensure that fungal infection and mycotoxin formation is as low as possible given the circumstances (CAC 2003; EC 2006). To enhance application of these practices and for further mycotoxin control, early insight into the expected DON contamination at harvest is useful to the various stakeholders

For Peer Review Only

of the wheat supply chain. As agronomical and climatic factors are assumed to affect mycotoxin formation during the entire growing season, prediction accuracy of DON in mature wheat is expected to increase with the progressing cultivation period. However, when developing predictive models, the needs of the different stakeholders of the wheat chain regarding the intended use of the predictions and their timing should be considered. Farmers may use DON predictions around wheat flowering to underpin decision making on the application of fungicides, using the most accurate predictions available at that critical growing stage. Feed and food industry, including collectors, millers, and processors, may use DON predictions to focus their inspections and to decide on buying, routing and processing of batches. They may need the DON prediction as early as available in the season and up to harvest, considering that the prediction accuracy will increase as the growing season progresses. Food safety authorities are expected to improve inspection efficiency when using mycotoxin predictions in constructing their yearly monitoring plan.

Worldwide, several initiatives are going on to develop quantitative predictive models for DON levels in mature wheat, however, only a few have been published (Schaafsma and Hooker 2007; Prandini et al. 2009). The most well-known model is DONcast, developed in Canada, aimed to assist farmers in their decision to spray (Hooker et al. 2002). The model predicts DON levels at wheat heading, and is based on regression analyses of field data. The model is commercialized to provide region specific DON predictions, using input on climate and agronomy (Schaafsma and Hooker 2007). Rossi et al. (2003) developed a mechanistic model, using experimental data, for (fungal infection and) accumulation of DON and zearalenole in wheat until harvest. The model is linked into a semi-quantitative system to estimate the risk for both mycotoxins (Rossi et al. 2007). A neural network model has been developed to

For Peer Review Only

predict DON content in mature wheat around flowering, using data from two four-year field experiments in Czech. The best fitting model includes preceding crop and four climatic related variables (Klem et al. 2007). For The Netherlands, a quantitative model to predict DON content in wheat at harvest has been developed recently (Franz et al. 2009). Model development was based on regression analyses using six years of field data. The best performing model included several agronomical variables and various climatic variables up to 24 days after wheat heading. This model could be useful in practice for feed and food industry, and food safety authorities as – in contrary to farmers - these stakeholders do not necessarily need the model results around wheat heading. For disease management, the postheading period is too wide, as hourly meteorological forecasts may be reliable up to 7 to 10 days, but loose reliability after this period.

The main aim of the current study is to develop quantitative models to predict DON levels in mature winter wheat in The Netherlands for different groups of end-users; more specifically, a model for farmers focusing on the wheat flowering period, and a model for industry and authorities including the period up to harvest. The central hypothesis of this study is that considering weather data during the entire period of cultivation will increase the prediction accuracy.

Material and Methods

Data collection

In the years 2001 – 2008, data were collected from in total 425 fields (records) throughout the Netherlands. The number of fields per year were: 29 in 2001, 48 in

For Peer Review Only

2002, 24 in 2003, 50 in 2004, 96 in 2005, 28 in 2006, 36 in 2007, and 114 in 2008. Information collected from these fields included: field location (zip code), wheat variety (VAR), cultivar resistance level (RES, levels 1-10), application of Fusarium fungicides (SPRAY, levels 0, 1, 2 times), and DON content (in μg kg^{-1}) in samples of mature wheat. Spikes were collected several days before harvest at random throughout the field until 1 kg was obtained. The entire grain sample was sent to an independent private laboratory in The Netherlands. DON was extracted and quantified using a liquid-chromatography with double in-line mass spectrometry (LC-MS/MS) (Spanjer et al. 2008), using a limit of quantification of 5 μg kg^{-1}. From the nearest meteorological station to each field, hourly weather data on temperature, rainfall and relative humidity was obtained for the entire observation period. Data from in total 29 meteorological stations were used. Part of the collected data was also used by Franz et al. (2009).

Data processing

In case of a sample not exceeding the limit of quantification, the DON level of the particular sample was set at this limit (of 5 μg kg^{-1}). The locations of the fields were divided into four regions, including North (N), West (W), South-East (SE) and North-West (NW), based on provinces (n=13) of The Netherlands, such that the four regions included a nearly equal number of provinces. For each record, the intensity of wheat production around the field, named "cereality" (CER), was calculated as the area of wheat grown divided by the total area. Field CER (in percentage per km^2, 0-100%) was calculated using data from the Dutch Ministry of Agriculture, Nature and Food Quality. For the wheat fields, flowering dates were registered by the farmers. However, some data were missing and observations showed to not to be precise in all

For Peer Review Only

cases. Therefore, for each record, the flowering date (FD) - representing the middle of the critical wheat flowering period - was calculated as the day at which the sum of average daily temperatures since January 1 of the particular year reached 1265˚C. Also, for each record, the harvest date (HD) was calculated as the first date at which the sum of the average daily temperatures after the (calculated) FD reached 989˚C (Lazar and Genovese 2004). Both FD and HD are expressed as the number of days defined from January 1 of the particular year. From the calculated FD and HD, the length of the period (in number of days) between flowering and harvest (leFH, in days) was derived.

For the development of the two models, i.e., the model for the farmers (model A) and the model for industry and authorities (model B), climatic variables were constructed in five relevant time windows of the wheat growing season. For model A, these time windows run up to 10 days after FD, whereas for model B they run up to harvest. The time blocks were defined on a daily basis with days running from 12 AM the previous day to 11 AM the actual day. Time blocks used in model A - represented relative to FD (day 0) - were: day -17 to -10 (A0), day -10 to -3 (A1), day -3 to 0 (A2), day 0 to 3 (A3), day 3 to 10 (A4). For model B, the time blocks included day -17 to -10 (B0), day -10 to 0 (B1), day 0 to 10 (B2), day 10 to 10 days before HD (B3), 10 days before HD to HD (B4). Note, that the number of days in time block B3 may vary between the records. For each of these time blocks, the following climatic variables were calculated per record:

- average hourly temperature (Tavg in ˚C)
- number of hours that temperature is 25˚C or higher (Th25)
- sum of hourly rainfall (rain in mm)
- total number of hours that relative humidity (RH) is 80% or higher (RHh80)

http://mc.manuscriptcentral.com/tfac Email: fac@tandf.co.uk
For Peer Review Only

- sum of hourly temperatures over the hours that RH is 85% or higher, provided this is the case for at least 2 consecutive hours (ThRH85_2)
- sum of hourly temperatures over the hours that RH is 85% or higher, provided this is the case for at least 4 consecutive hours (ThRH85_4)

Model development and validation

In model development, a variety of statistical analyses was performed, including: descriptive statistics, and univariate and multiple regression analyses. All statistics were performed using GenStat 11th Edition (GenStat Committee, Algorithm group Inc., see http://www.vsni.co.uk/), using a significance level of 0.05. Given the range of DON values (the response variable) and the expected increasing variation at higher levels of DON, it was assumed that DON follows a lognormal distribution. Therefore, all statistical analyses were performed on logtransformed data, i.e., the natural logarithm of DON (lnDON).

Univariate regression analyses were run to test the relationships of the individual agronomical variables, being REG, VAR, RES, SPRAY, CER, FD and leFH, with lnDON. The variables that contributed significantly in explaining variation in the DON values were considered in further analyses. Pearson correlation coefficients were calculated for all combinations of the various climatic variables in the different time blocks. Climatic variables that were sufficiently uncorrelated ($r<0.7$) were taken into account in the further selection process. Effects of quadratic terms of the climatic variables in the different time periods and of relevant interactions between climatic variables were also tested univariately. The resulting agronomical and climatic variables, with relevant quadratic terms and interactions, were considered in the multiple regression analyses ('all possible subset selection'

For Peer Review Only

procedure using GenStat RSEARCH). The best set of variables was identified by evaluating the percentage variance accounted for (R^2_{adj}) and the value of *Mallows'* C_p (Archdeacon 1994; Gross 2003). Models with the highest R^2_{adj} and a low *Mallows'* C_p that included only significant variables were selected. All the above described steps for the selection of variables were done for constructing model A and B separately, and using the entire dataset. However, three records that showed to result in a very high leverage trough all the modeling steps were not considered in the analyses.

Once model A and B were selected, they were fitted to 90% of the records, and validated using the remaining 10% of the records. For each model, the two data subsets were constructed by drawing a random sample of 10% from each year-region combination from the entire dataset (thus resulting in a 10% subset for model A and a 10% subset for model B). The two 90% data subsets were checked for balanced number of records and DON levels over years, regions, cultivar and spraying levels. Parameter estimates and standard errors (s.e.) were obtained by fitting the final models to their respective 90% data subsets. These final models were checked for data with high leverages, correlations between model variables, residuals (plots) and figures of the fitted and observed relationships between the individual model parameters and lnDON. Based on the two final models, predictions (with s.e.) of DON levels were calculated for the 10% data subsets. The correlation coefficients between these predicted values and the corresponding observed DON levels were calculated. Also, the 95% confidence intervals of the predictions were calculated.

Results

For Peer Review Only

Descriptive analyses

The overall average observed DON level (min – max value, n records) in the period 2001-2008 was 491.8 µg kg^{-1} (5.0 – 10400, n=425). Average DON levels varied highly between years, from a lowest value of 5.0 µg kg^{-1} in 2006 to a highest value of 2045.3 µg kg^{-1} in 2007 (Table 1). Yearly average values in 2001, 2002, 2004 and 2008 were around the overall average. In 2006, all samples were found to be below the limit of quantification.

The distribution of observations over regions showed a comparable number of records in three regions (M, N, W), but a lower number in the S-region. Observed average DON levels in mature winter wheat were comparable for the regions M, W and S, being 424.5 µg kg^{-1} (5.0 – 4600, n=170), 357.2 µg kg^{-1} (5.0 – 3150, n=106) and 572.3 µg kg^{-1} (5.0 – 1300, n=23), respectively. DON levels were somewhat higher in the N-region, averaging 681.0 µg kg^{-1} (5.0 - 10400, n=126). DON levels were lower in case of spraying one or two times, as compared to no fungicide treatment; averages were 299.3 µg kg^{-1} (22.6 – 1300, n=66) for two applications, 291.9 µg kg^{-1} (5.0 – 3217, n=91) for one application, and 607.0 µg kg^{-1} (5.0 – 10400, n=268) in case no fungicide was applied. In total 27 different wheat varieties were used, with 13 of them having less than 5 observations. Wheat resistance level (RES) ranged from 5.0 to 8.5 with a median of 6.5. Average DON values run from 682.5 µg kg^{-1} for a RES of 5.0, around 500 µg kg^{-1} for a RES of 6.5-7.5 (468.2 µg kg^{-1} and 507.7 µg kg^{-1} for a RES of 6.5 and 7.5, respectively), to a lowest value of 137.5 µg kg^{-1} for a RES of 8.5. CER ranged from 1.6% to 62.5% at the highest, with an average of 17.8%. FD ranged from 119 at the lowest to 173 at the highest, it's median being 155. LeFH ranged from 30 to 76 days with a median value of 56. Observed DON content in mature wheat increased with a higher CER, an earlier FD, and an increased LeFH.

For Peer Review Only

Regression analyses

Univariate analyses that showed year had a highly significant effect on DON levels, and solely accounted for 52.8 % of the variance (s.e. 1.02). However, year was excluded as a potential explanatory variable in further analyses as it was assumed its effects are due to the effects of other variables. Additionally, the variable year is not useful for making predictions for the future. Individual effects of REG, VAR, RES, SPRAY, FD, and leFH on DON levels also were significant and accounted for, respectively, 3.3%, 4.9%, 2.4%, 1.1%, 20.5%, and 30.1 % of the variation. The (slight) effect of CER was not significant. The variables VAR and RES are (by definition) related to each other, and given the low number of observations in some VAR classes, this variable was excluded from further model building. Based on these results, agronomical variables included in constructing both models were: REG, RES, SPRAY, FD, added with leFH for model B.

Model A (flowering period) Within each of the five A time periods, the correlation between the variables RHh80, ThRH85_2 and ThRH85_4 showed to be very high. Out of these three variables related to relative humidity, RHh80 is the simplest one to record and, therefore, this variable was chosen for further model building. The variables Tavg and Th25 also were strongly correlated with each other; correlation coefficient values ranged from 0.77 to 0.86 in the different A time blocks. Correlations between the same climatic variables in different A time blocks in general showed to be low. The highest correlation was found for Tavg in A2 and Tavg in A3 (r= 0.67); all other correlations were below 0.60.

For Peer Review Only

For model A, the best fitting multiple regression model included the following variables: REG, RES, SPRAY, FD, Tavg (with quadratic effects) in all A time blocks except for A2, RHh80 in A0 and A3, rain in A1 and A4 (with quadratic effect), and the interaction between rain in A1 and Tavg in A3. The results of fitting this model to its 90% data subset are presented in Table 2. All variables showed to be significant ($p<0.05$). The percentage of variance accounted (R^2) for was 64.4%, the standard error of observations was 0.87.

Model B (entire cultivation period) In all B time periods, the correlation between ThRH85_2 and ThRH85_4 showed to be very high. This also accounted for each of these two variables and RHh80, except for the period B4. The correlation between Tavg and Th25 in individual B time periods was also high, varying from 0.77 to 0.96. Correlations between the same climatic variables in different B time blocks were low; the highest value was found for Tavg in B0 and Tavg in B4, equalling 0.68. The interaction between rain in B1 and Tavg in B2 showed significantly to affect DON levels univariately.

For model B, the best fitting set of variables included: REG, RES, SPRAY, FD, leFH, rainfall and RHh80 in B0, and Tavg in B0, B1, B3 and B4 (with quadratic effects in B3 and B4). The results of fitting this model to its 90% data subset are presented in Table 3. The percentage variance accounted for (R^2) was 65.6%, the standard error of observations was 0.836. The correlation between leFH and Tavg in B3 was somewhat high ($r= -0.72$). However, both variables had a significant effect in the final model, and were included based on their logical relationship.

Model validation

http://mc.manuscriptcentral.com/tfac Email: fac@tandf.co.uk
For Peer Review Only

Fitting the final models A and B to their respective 10% data subsets resulted in a correlation coefficient between predicted and observed DON levels of 0.84 and 0.78, respectively (n=41 records). The relationships between the predicted and observed DON levels are illustrated in Figures 1 and 2 for model A and B, respectively. All observed DON values were within the 95% confidence interval of the predicted values.

For Peer Review Only

Discussion

The two models constructed in the current study present the second attempt to predict DON levels in mature winter wheat in The Netherlands. Compared to the (basic) model of Franz et al. (2009), the two current models are more specified to the intended user-groups, particularly wheat growers (model A) and other stakeholders, including feed and food industry and food safety authorities (model B). Model A includes climatic variables in a shorter period around wheat flowering, i.e., up to 10 days after FD, as compared to the basic model that used the period up to 24 days after wheat heading (Franz et al. 2009). Unless this shorter period, the percentage of explained variance was higher (R^2 of 65% vs 59%). The agronomical variables in both models are comparable, as are the climatic variables, all being related to rainfall, temperature and relative humidity in the pre- and postflowering periods. The increased predictive power of model A of this study might have been due to the additional climatic variables considered - in particular the inclusion of quadratic effects and interactions - and/or the different distinction between days. In this study, days were considered to run from 12 AM the previous day to 11 AM the actual day such to include potential overnight effects of climate, specifically the effect of

extended periods of leaf wetness (dew) during the night and early morning on fungal development.

This study showed that taking into account the entire period up to harvest (model B), as compared to flowering period (model A), increased the accuracy of the model predictions; the percentage explained variance of model B (R^2=65.6%) was slightly higher, as compared to this metric for model A (R^2=64.4%), and the standard error of observations was somewhat lower. The agronomical variables in models A and B are identical, however, model B also includes the length of the period from FD to HD. The longer this period the higher the DON levels at harvest. This effect did not dependent on effects of the actual FD and wheat resistance level. Given the negative relationship between FD and DON values, the (positive) effect of an increasing length from FD to HD may be related to extended conducive periods of humidity and high temperatures during mid summer.

Both in model A and B, climatic variables in the preflowering period relate to average temperature, rainfall and relative humidity. Both a higher rainfall and an increase in the number of hours that RH is 80% or more increase DON levels at harvest. A higher average temperature in the early preflowering period (-17 to -10 days FD) increases DON levels, but a higher average just before FD decreases DON levels. The latter finding may be related to the humid conditions that often go hand in hand with low temperatures and are conducive for fungal infection in the preflowering period. Climatic effects in the postflowering period relate to rainfall, temperature and relative humidity in case a 10 day period is considered (model A). However, temperature is the only significant factor when the entire postflowering period is taken into account (model B). These findings of climatic effects during the postflowering period correspond well to Franz et al. (2009). In model A, an interactive effect of the

For Peer Review Only

preflowering rainfall and postflowering temperature on DON at harvest was found, with a more profound effect of increasing temperature after FD (i.e., higher DON levels) in case of a higher rainfall before FD. In the two current models, the majority of the temperature related variables in the different periods of wheat growing showed to have an optimal effect on DON levels at about 16-20 ˚C. The effects of the climatic variables in the two models of the current study reflect the ecological conditions of infection of *Fusarium* spp. in wheat and consequent DON formation, with increasing effects of moist (wet and humid) and warm conditions around flowering (Osborne and Stein 2007).

In the construction of both models, the intensity of wheat production in the area around the wheat field (cereality) was considered in attempt to unravel the region effect on DON values, found in the current study and previously (Hooker et al. 2002; Franz et al. 2009). It was hypothesized that the more wheat produced in the area surrounding the wheat field, the higher the *Fusarium* spp. infection pressure. However, cereality did not significantly affect DON levels at harvest and, obviously, the region effect is caused by other factors. These might include unanalyzed climatic factors such as precrop winter conditions, regional variations in farm management and/or in the genetic composition of *Fusarium* spp. These potential factors underlying the region effect should be further investigated.

Several agronomical variables that are previously reported to affect DON contamination in mature winter wheat, particularly tillage practices and previous crop (Champeil et al. 2004a; Schaafsma et al. 2005; Beyer et al. 2006; Eiblmeier et al. 2007), were not considered in model construction, primarily because data on these two factors were not available. Secondly, variability in these two factors was (expected to be) small, making them not useful for model construction. In The

For Peer Review Only

Netherlands, generally, deep ploughing is applied and maize is not grown before wheat (unpublished data), hereby avoiding the conducive effects of growing grain maize as preceding crop and applying minimum or not tillage on DON contamination. Would the models be applied to countries with other tillage and cropping practices, it is worthwhile to investigate the possible benefits of considering these variables in model construction, at least for the farmer's model. Industry and authorities often do not have data on precrop and tillage practices and therefore, may prefer one of the current models.

Corresponding to the two comparable previous modeling studies (Hooker et al. 2002; Franz et al. 2009), the current study showed that DON levels in mature winter wheat can be predicted accurately using regression analyses of field data, and considering a narrow time window around wheat flowering. The predictive performance of models of the three studies, expressed by the percentage explained variation (R^2), is in the same order of magnitude, being 73% for DONcast in Canada (Hooker et al. 2002), 64.4% for model A of this study, and 59% for the previous Dutch model (Franz et al. 2009). The r-square reported for DONcast (Hooker et al. 2002; Hooker et al. 2004; Schaafsma and Hooker, 2007) is somewhat higher than this figure for the three Dutch models. This might – at least partly – be related to the more uniform conditions of wheat cultivation in Canada (climate, field size etc.) as compared to The Netherlands. A previous study showed that application of DONcast to the Dutch field data resulted in very poor performance (Franz et al., 2009), justifying the development of models for this country. Taking into account the entire cultivation period (model B) as compared to the flowering period (model A) increased the accuracy of the DON predictions for The Netherlands, albeit this increase is rather small. Therefore, the intended end-users of model B (feed and food industry and/or

For Peer Review Only

authorities) might as well prefer to use model A, providing the DON predictions much earlier in the season. The costs and benefits of the increased prediction accuracy against the additional time could be further investigated by application of both models into practice.

The construction of predictive models for mycotoxins in wheat and their statistical performance not only depends on the specific variables considered in the model building process but also on the quality, particularly the precision, of the underlying data used, e.g., the extent to which the nearest meteorological station represents the climate of the particular wheat field. In addition, the wheat samples and measured DON levels are known to be subject to effects of sampling error and characteristics of the analytical method (e.g., sensitivity, specitivity, detection limit) used (Josephs et al. 2001; Whitaker et al. 2002; Whitaker 2003; Champeil et al. 2004b). At least part of the model residues is due to the variation in the response variable (DON levels) and the explanatory (agronomical and climatic) variables caused by one or more of these effects. Improving the performance of the predictive models could, therefore, possibly be obtained by using data of even higher quality.

The input variables used in constructing the current models represents a combination of the underlying process of fungal infection of the wheat and mycotoxin formation in the different critical periods of wheat growing. Expertise on the ecology of these processes was used to construct the variables and the time blocks in model development. When increased insight into the factors affecting the behavior of *Fusarium* spp and mycotoxin formation becomes available it should be included in the modeling process to further improve the definition of possible input variables and variable selection in model construction. Also, feedback from the end-users, applying

For Peer Review Only

the predictive models into practice, to model developers could improve model development.

Conclusions

This study showed that DON levels in mature winter wheat can be predicted accurately by regression models that are constructed using field data on agronomy and climate during the wheat growing season. Both the model for the farmer and the model for industry/authority include region, spraying, wheat resistance level, flowering date, various climatic variables related to rainfall, relative humidity and temperature in different critical periods of wheat cultivation, as well as the length between flowering and harvest (model for industry/authority only). DON levels at harvest can be predicted accurately around wheat flowering. Taking into account the entire period up to harvest slightly improves the prediction accuracy. The two models are expected to be useful for the various stakeholders of the wheat supply chain in order to control DON contamination of wheat derived feed and food products.

Acknowledgements

The research presented in this paper was financed by the Dutch Product and Food Safety Authority (VWA) and the Dutch Product Board for Arable Products (HPA). The authors greatly acknowledge all persons involved in data collection: Harro Spits, Huub Schepers, Don Jansen, Cees Waalwijk, Plant Sciences Group (Wageningen UR). Paulien Adamse and Eelco Franz, RIKILT – Institute of Food Safety and Michel Uiterwijk, Alterra (Wageningn UR) are kindly thanked for assistance in data

For Peer Review Only

processing.

For Peer Review Only

References

Archdeacon TJ. 1994. Correlation and regression analysis: a historian's guide. Madison, Wisconsin, USA: The University of Wisconsin Press, 352 pp.

Beyer M, Klix MB, Klink H, Verreet J-A. 2006. Quantifying the effects of previous crop, tillage, cultivar and triazole fungicides on the deoxynivalenol content of wheat grain - A review. Journal of Plant Diseases and Protection 113:241-246.

Bottalico A, Perrone G. 2002. Toxigenic *Fusarium* species and Mycotoxins Associated with Head Blight in Small-Grain Cereals in Europe. European Journal of Plant Pathology 108:611–624.

Champeil A, Doré T, Fourbet J-F. 2004a. Fusarium head blight: epidemiological origin of the effects of cultural practices on head blight attacks and the production of mycotoxins by Fusarium in wheat grains. Plant Science 166:1389-1415.

Champeil A, Fourbet J-F, Doré T. 2004b. Effects of grain sampling procedures on Fusarium mycotoxin assays in wheat grains. Journal of Agricultural and Food Chemistry 52:6049-6054.

Codex Alimentarius Commission (CAC). 2003. Code of practice for the prevention and reduction of mycotoxin contamination in cereals, including annexes on ochtratoxin A, zearalenone, fumonisins and tricothecences. CAC/RCP 51-2003, 8 pp. Available at: http://www.codexalimentarius.net/download/standards/406/CXC_051e.pdf. Accessed 24 September 2009.

Creppy EE. 2002. Update of survey, regulation and toxic effects of mycotoxins in Europe. Toxicology Letter 127:19-28.

For Peer Review Only

Diekman MA, Green ML. 1992. Mycotoxins and reproduction in domestic livestock. Journal Animal Science 70:1615-1627.

D'Mello JPF, Placinta CM, Macdonald AMC. 1999. Fusarium mycotoxins: a review of global implications for animal health, welfare and productivity. Animal and Feed Science Technology 80:183-205.

Eiblmeier P, Von Lepschy Gleissenthall J. 2007. Risk evaluation of deoxynivalenol levels in Bavarian wheat from survey data. Journal of Plant Diseases and Protection 114:69-75.

European Commission (EC). 2006. Commission Recommendation of 17 August 2006 on the prevention and reduction of Fusarium toxins in cereals and cereal products (2006/583/EC). Official Journal of the European Union L234:35-40.

Franz E, Booij CJH, Van der Fels-Klerx HJ. 2009. Prediction of Deoxynivalenol content in Dutch winter wheat. Journal of Food Protection 72: 2170–2177.

Gross J. 2003. Lecture Notes in Statistics 175 - Linear Regression. 2003. Berlin, Germany: Springer-Verlag, 394 pp.

Hooker DC, Schaafsma AW, Tamburic-Ilincic L. 2002. Using weather variables pre- and post-heading to predict deoxynivalenol content in winter wheat. Plant Disease 86:611-619.

Hooker DC, Schaafsma AW. 2004. The DONcast model: predicting deoxynivalenol (DON) in wheat. In: Canty SM, Boring T, Wardwell J and RW Ward (Eds.), Proceedings of the Second International Symposium on Fusarium Head Blight; incorporating the Eight European Fusarium Seminar, held 11-15 December 2004. Orlando, FL, USA. East Lansing, MI, Michigan State University, vol. 2, p. 458.

For Peer Review Only

Josephs, RD, Schuhmacher R, Krska R. 2001. International interlaboratory study for the determination of he *Fusarium* mycotoxins zearalenone and deoxynivalenol in agricultural commodities. Food Additives and Contaminants 18:417-430.

Klem K, Váňová M, Hajšlová J, Lancová K, Sehnalová M. 2007. A neural network model for prediction of deoxynivalenol content in wheat grain based on weather data and preceding crop. Plant Soil Environment 53:421–429.

Lazar C, Genovese G. 2004. Methodology of the MARS crop yield forecasting system. Volume 2. Agrometeorological data collection, processing and analysis. European Communities (ISBN 92-894-8181-1).

Miller JD. 2008. Mycotoxins in small grains and maize: Old problems, new challenges'. Food Additives and Contaminants 25:219-230

Minervini F, Aquila MED. 2008. Zearalenone and reproductive function in farm animals. International Journal of Molecular Sciences 9:2570-2584.

Morgavi DP, Riley RT. 2007. An historical overview of field disease outbreaks known or suspected to be caused by consumption of feeds contaminated with Fusarium toxins. Animal and Feed Science Technology 137:201-212.

Osborne LE, Stein JM. 2007. Epidemiology of Fusarium head blight on small-grain cereals. International Journal of Food Microbiology 119:103-108.

Prandini A, Sigolo S, Filippi L, Battilani P, Piva G. 2009. Review of predictive models for *Fusarium* head blight and related mycotoxin contamination in wheat. Food and Chemical Toxicology 47:927-931.

Rossi V, Giosuè S, Pattori E, Spanna F, Del Vecchio A. 2003. A model estimating the risk of Fusarium head blight on wheat. OEPP/EPPO Bulletin 33:421–425.

Rossi V, Giosuè S, Terzi V, Scudellari D. 2007. A decision support system for Fusarium head blight on small grain cereals. OEPP/EPPO Bulletin 37:359-367

For Peer Review Only

Schaafsma AW, Tamburic-Ilincic L, Hooker, DC. 2005. Effect of previous crop, tillage, field size, adjacent crop, and sampling direction on airborne propagules of Gibberella zeae/Fusarium graminearum, fusarium head blight severity, and deoxynivalenol accumulation in winter wheat. Canadian Journal of Plant Pathology 27:217-224.

Schaafsma AW, Hooker DC. 2007. Climatic models to predict occurrence of *Fusarium* toxins in wheat and maize. International Journal of Food Microbiology 119:116-125.

Smith JE, Solomons G, Lewis C, Anderson JG. 2006. Role of mycotoxins in human and animal nutrition and health. Natural Toxins 3:18-192.

Spanjer MC, Rensen PM, Scholten JM. 2008. LC-MS/MS multi-method for mycotoxins after single extraction, with validation data for peanut, pistachio, wheat, maize, cornflakes, raisins and figs. Food Additives and Contaminants 25:472-489.

Whitaker TB, Hagler Jr WM, Giesbrecht FG, Johansson AS. 2002. Sampling wheat for deoxynivalenol**.** Advances in Experimental Medicine and Biology 504:73-83.

Whitaker TB. 2003. Standardisation of mycotoxin sampling procedures: An urgent necessity. Food Control 14:233-237.

For Peer Review Only

Table 1. Observed deoxynivalenol levels (in µg kg^{-1}) in mature winter wheat in The Netherlands in the period 2001-2008

	2001	2002	2003	2004	2005	2006	2007	2008	overall
Mean	648.8	747.3	193.9	323.3	244.0	5.0	2045.3	318.5	491.8
Median	111.2	297.1	135.3	276.2	130.0	5.0	1250.0	150.0	170.0
Min	22.60	34.70	99.00	40.0	30.0	5.0	140.0	50.0	5.0
Max	3495	3788	464	1210	850	5.0	10400	4600	10400
Lower Quantile	73.1	95.3	99.0	140.0	75.0	5.0	360.0	77.0	81.3
Upper Quantile	733.2	922.2	276.8	420.0	371.2	5.0	3025.0	340.0	434.4
N	29	48	24	50	96	28	36	114	425

For Peer Review Only

Table 2. Summary of the final model A for prediction of DON levels in mature winter wheat in The Netherlands[1]

Model variable[2]	Parameter estimate[2]	Standard error	Parameter significance
Constant	-16.04	7.62	0.036
REG N	-0.337	0.142	0.018
REG W	-1.654	0.167	<0.001
REG S	-0.812	0.312	<0.010
SPRAY 1	0.502	0.188	0.008
SPRAY 0	0.720	0.175	<0.001
Flowering date	-0.12343	0.00865	<0.001
Resistance level	-0.4031	0.0484	<0.001
Rain A1	-0.1362	0.0311	<0.001
Tavg A3	-0.0414	0.0424	0.329
Interactive effect rain A1 * Tavg A3	0.00911	0.00211	<0.001
Rain linear A4	-0.0977	0.0120	<0.001
Rain quadratic A4	0.001501	0.000196	<0.001
Tavg linear A0	1.165	0.384	0.003
Tavg quadratic A0	-0.0315	0.0127	0.013
Tavg linear A1	0.779	0.364	0.033
Tavg quadratic A1	-0.0304	0.0108	0.005
Tavg linear A4	3.494	0.673	<0.001
Tavg quadratic A4	-0.1124	0.0200	<0.001
RHh80 A0	0.02194	0.00308	<0.001
RHh80 A3	0.01493	0.00485	0.002

[1]The model predicts ln(DON) concentrations (in µg kg^{-1}). DON values (in µg kg^{-1}) can be obtained by back transformation: $e^{(\text{model A})}$.

For Peer Review Only

[2]Model variables are defined in the body text of this manuscript. Parameter estimates for the factors REG and SPRAY are differences compared with their reference levels, being REG M and SPRAY 0, respectively.

For Peer Review Only

Table 3. Summary of the final model B for prediction of DON levels in mature winter wheat in The Netherlands[1]

Model variable[2]	Parameters estimate[2]	Standard error	Parameter significance
Constant	-31.39	9.03	<0.001
REG N	0.246	0.194	0.206
REG W	-0.729	0.159	<0.001
REG S	0.305	0.260	0.241
SPRAY 1	0.688	0.165	<0.001
SPRAY 0	0.842	0.155	<0.001
Flowering date	-0.0647	0.0102	<0.001
Resistance level	-0.4281	0.0463	<0.001
Length FH	0.0888	0.0173	<0.001
Rain B0	0.03869	0.00572	<0.001
RHh80 B0	0.00633	0.00252	0.012
Tavg B0	0.1076	0.0477	0.025
Tavg B1	-0.0654	0.0402	0.104
Tavg linear B3	2.614	0.965	0.007
Tavg quadratic B3	-0.0781	0.0271	0.004
Tavg linear B4	2.18	1.02	0.033
Tavg quadratic B4	-0.0570	0.0277	0.040

[1]The model predicts ln(DON) concentrations (in µg kg^{-1}). DON values (in µg kg^{-1}) can be obtained by back transformation using $e^{(\text{model B})}$.

[2]Model variables are defined in the body text of this manuscript. Parameter estimates for the factors REG and SPRAY are differences compared with their reference levels, being REG M and SPRAY 0, respectively.

For Peer Review Only

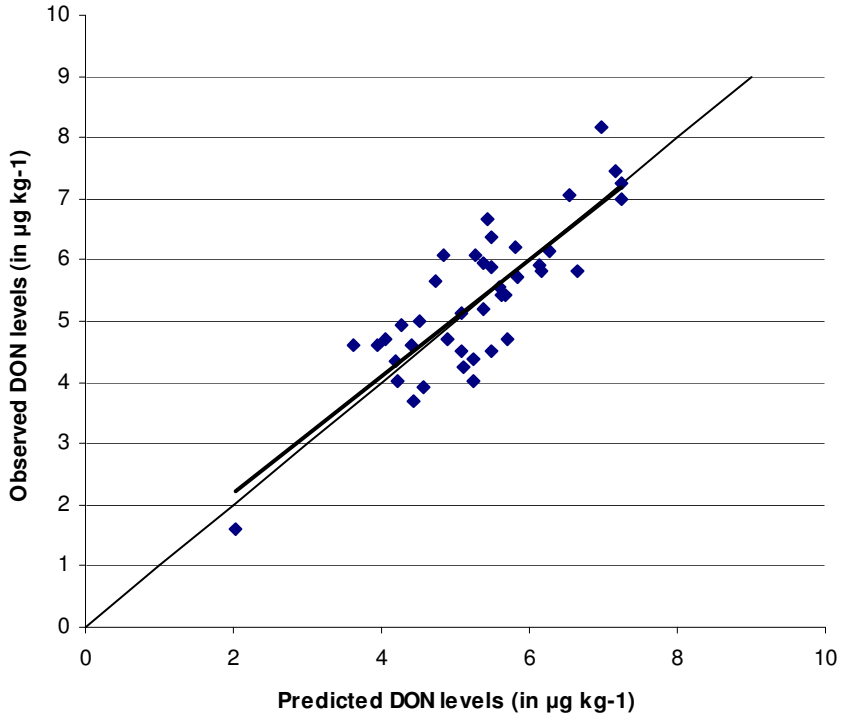
FIGURE CAPTIONS

Figure 1. Predicted DON levels using model A vs observed DON levels (in μg kg^{-1}) in mature winter wheat in The Netherlands

Figure 2. Predicted DON levels using model B vs observed DON levels (in μg kg^{-1}) in mature winter wheat in The Netherlands

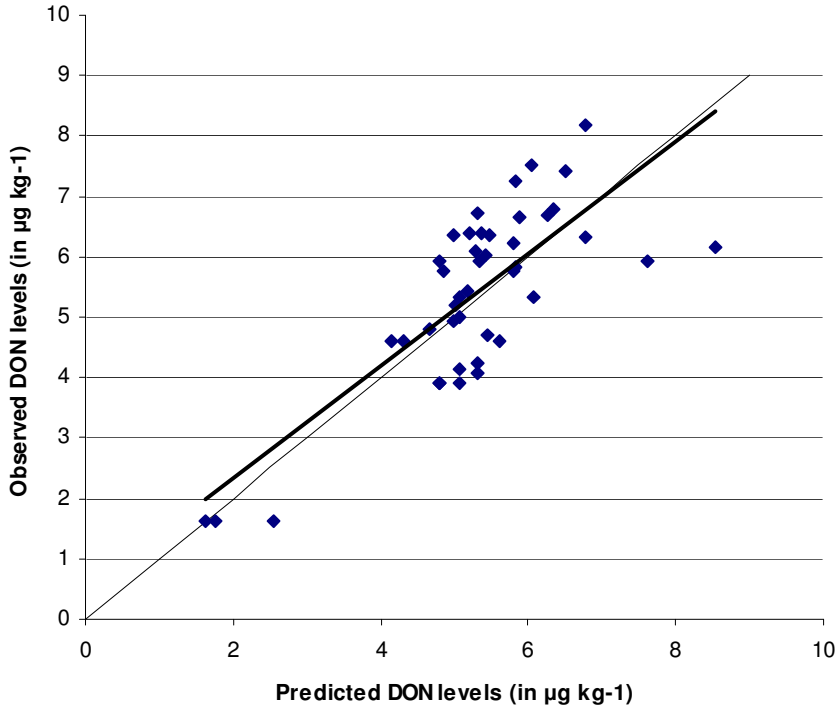
For Peer Review Only

Figure 1.



For Review Only

Figure 2.



Review Only