Modelling the propulsion system of a combine harvester

Tom Coen, Ivan Goethals, Jan Anthonis, Bart De Moor and Josse De Baerdemaeker

***Abstract*— The goal of this paper is to construct a model for the propulsion system of a combine harvester. Linear as well as nonlinear techniques are used. The influence of varying factors such as slope and mass present in the grain bin are examined. To conclude, the effect of the diesel rpm level and of the gear is included in the model.**

I. INTRODUCTION

Over the last decade combine harvesters have evolved to state of the art machinery [1] [2] [3] [4]. A lot of efforts have been made to alleviate the task of the combine operator. This is necessary and economical because of the high work load during harvest time.

A further improvement could be the introduction of cruise control to maintain constant speed under disturbances such as varying slope and mass. The combine harvester considered here is a New Holland CR combine. This paper discusses the identification of the steering current of the hydrostatic transmission towards the resulting machine speed.

First of all the goal of the model and the model specifications are presented. In a following step a model for the empty combine on a flat concrete road is identified. In order to derive the model, a white noise sequence is applied to the steering current.

The mass of a combine can vary between about 15 tonnes (empty) and 20 tonnes (fully loaded with grain). This mass difference may have a significant influence on the machine dynamics [5]. The influence of the instantaneous machine mass and the slope of the road are examined.

A. System overview

The propulsion system of a combine harvester consists of a diesel engine that operates on a constant rpm level. The rpm level of this engine is guarded by an independent controller. This diesel engine drives the hydrostatic pump. The flow of this pump is then controlled by a steering current.

The pump drives an hydrostatic engine, which in its turn is coupled to a mechanical transmission with 4 gears. Through a differential the wheels of the front axle are driven.

Because of special precautions, the machine is able to brake actively. This means that the hydrostatic pump starts functioning as a generator. Furthermore, the hydrostatic pump shows a static hysteresis. It is clear that this increases the system complexity significantly.

T. Coen, J. Anthonis and J. De Baerdemaeker are with the Faculty of Bio-Engineering, Laboratory for Agro-Machinery, Katholieke Universiteit Leuven, Kasteelpark Arenberg 30, 3001 Leuven, Belgium

I. Goethals and B. De Moor are with the Faculty of Applied Sciences, ESAT-SCD, Katholieke Universiteit Leuven, Kasteelpark Arenberg 10, 3001 Leuven, Belgium

tom.coen@agr.kuleuven.ac.be (corresponding author)

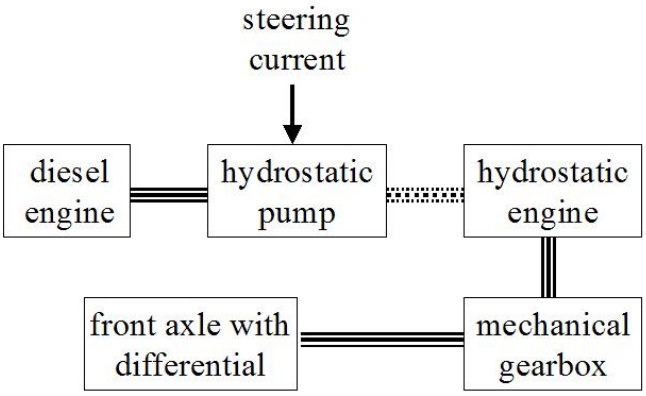


Fig. 1. Schematic overview of the propulsion system

B. Cruise control

The objective of cruise control is to maintain the machine speed on a user defined setpoint. This relationship depends on the machine mass, slope and soil type. It is clear that machine mass and slope influence the inertial force F_I and gravity force F_G applied on the machine. These forces can be written as:

$$F_I = m\dot{v} \quad (1)$$

$$F_G = mg\sin\alpha \quad (2)$$

with m the total machine mass, v the machine speed, g the gravitational acceleration and α the slope.

In an hydraulic system however, such load variations are for a large part absorbed by the hydraulic pump [6].

II. MODEL DESCRIPTION & SPECIFICATIONS

A. Model Description

The model will describe the relationship between input current and resulting ground speed. It will be derived using system identification.

The possible advantages of non-linear identification over more classical linear techniques will be examined.

B. Model Specifications

Before identifying a model, some questions have to be answered, such as:

- In which frequency range will the model be used?
- How far ahead does the model have to predict?
- At which sampling frequency will the model operate?

The first question leads to 2 other questions. Firstly, what is the highest frequency to which the machine is susceptible,

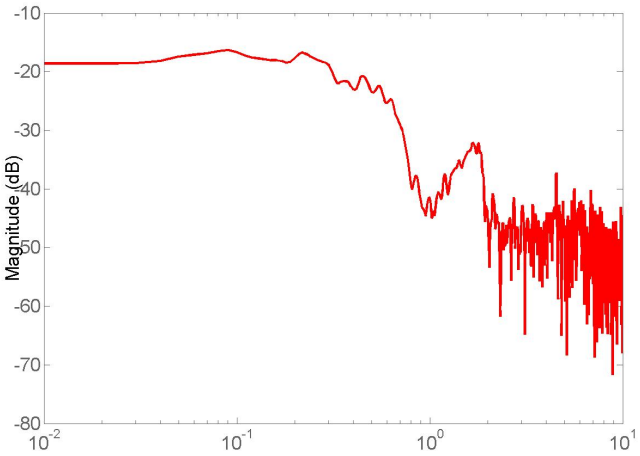


Fig. 2. The FRF between input current and ground speed

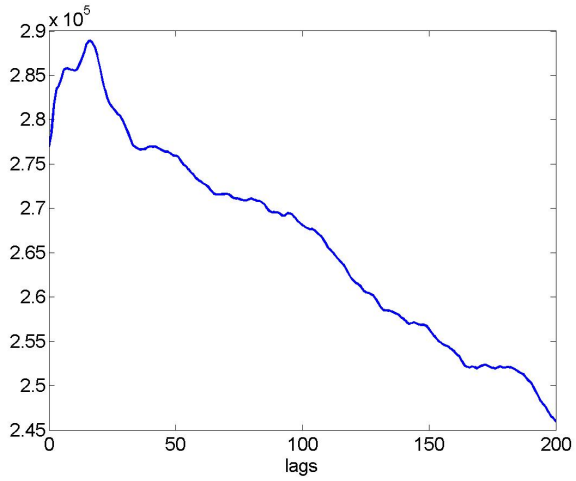


Fig. 3. Cross correlation between input current and ground speed

and secondly, what is the highest frequency at which the controller will operate?

In Fig. 2 the frequency spectrum of the machine is shown. The machine acts as a low pass filter with cut-off at $1Hz$. Between $1Hz$ and $2Hz$ there is a bump in the FRF visible, probably due to nonlinear effects. As far as the second question is concerned, the controller will not be allowed to use higher frequencies than $1Hz$ in order to guarantee the comfort of the operator. The model thus needs to cover frequencies up to $1Hz$.

In the field the machine speed is about 1 m/s. Looking 3 seconds ahead thus implies looking 3 meters ahead. Taking the inertia of the machine into account, a prediction horizon of 3 seconds suffices. All model validations are performed by predicting a sequence 3 seconds ahead.

The software of the production version of the harvester operates at frequencies between $5Hz$ and $100Hz$. The sampling frequency should be chosen in function of the desired frequency range. The Nyquist theorem links the minimum sampling frequency to the frequency content. Out of calculation time considerations the sampling frequency should be lowered as much as possible. Models at $20Hz$ and $5Hz$ are derived.

III. IDENTIFICATION EXPERIMENTS

The data used in this paper was obtained from excitations on a New Holland CR combine harvester. All the input and output signals are sampled at $20Hz$. This is sufficiently fast to capture all the relevant machine dynamics. The model identification is done for 2^{nd} gear since this is the normal harvesting gear.

For the model identification white noise sequences were applied to the input current of the hydrostatic pump. These sequences are very well suited for identification. Experiments are performed with different machine masses (15.700 kg, 17.200 kg and 18.800 kg), and on flat terrain, slopes uphill and slopes downhill (of about 3^o).

Step inputs are also applied in these different cases, since these are the most common inputs given by the operator. On flat terrain these step inputs are also applied in the other gears to examine the influence of the different gears.

IV. SIMPLE ROAD MODEL

A. Data exploration

A first step is to identify the delay between input current and ground speed. This can be done by calculating the cross correlation function between input current and ground speed for a white noise sequence (Fig. 3). The maximum of the cross correlation is about at a 15 samples lag, which implies a delay of 0.75 seconds.

B. Linear models

For the linear models 2 sets of parameters need to be tuned. On one hand the orders of the system need to be chosen, on the other hand the delay from input to output needs to be fixed. It is clear that choosing the delay too small will imply the estimation of zeros, whilst choosing the delay too large will degrade the information contained in the data. Considering Fig. 3, a delay of about 0.5 seconds is good starting value.

The models are obtained with the System Identification Toolbox [7] supplied with Matlab. In this section models are then validated on an independent white noise sequence. This gives an idea of the dynamic performance of the models. Later on the performance on step inputs will be examined.

The input and output data is filtered by a 10th order Butterworth filter at $1Hz$. If the cut-off frequency is taken at for example $2Hz$, linear modelling becomes much harder. This can be understood in Fig. 2. The bump in the FRF between $1Hz$ and $2Hz$ severely complicates the modelling. The validation of the derived models is done on unfiltered output data. In this way is shown that the unmodelled output higher than $1Hz$ does not influence the model performance to a high degree.

TABLE I

RESULTS OF ARX WITH DIFFERENT ORDERS AND DELAYS, 3 SECONDS AHEAD PREDICTION

ID	Order A	Order B	Delay(s)	Perf. $20Hz$	Perf. $5Hz$
arx1	10	10	0	-	-
arx2	0	10	0	76.4%	77.6%
arx3	10	0	0	-	-
arx4	0	10	0.5	77.1%	78.2%
arx5	0	5	0.5	77.2%	78.0%
arx6	0	4	0.5	77.0%	77.8%
arx7	0	3	0.5	75.1%	76.4%
arx8	0	2	0.5	63.0%	64.1%
arx9	0	3	0	51.8%	61.6%
arx10	1	3	0.5	68.2%	72.3%

- indicates an unstable prediction model

The mean is subtracted from the training data before deriving the model.

1) ARX: The ARX model has the following form:

$$A(q)y(t) \quad = \quad B(q)u(t-d) + e(t) \tag{3}$$

where the order of A and B and the delay d in the input u are parameters.

As a first step some different orders are tried and evaluated. The data is then decimated to $5Hz$ sampling frequency because this is the lowest frequency used on production machinery. The models are evaluated on the percentage of variation in the data that is explained by the model (3 seconds ahead prediction).

Taking the model complexity into account, the best choice for the model order would be 3 in the input and 0 in the output with a delay of 0.5 seconds (see Table I). This means that only three parameters need to be estimated.

$arx10$ has almost white residuals, whilst this is not the case for the $arx7$. The best performing ARX models are $arx7$ and $arx10$.

2) ARMAX: An ARMAX model has the following form:

$$A(q)y(t) \quad = \quad B(q)u(t-d) + C(q)e(t) \tag{4}$$

The advantage of ARMAX over ARX thus lies in the higher flexibility of the noise term. However, starting from the ideal ARX parameters (being 0 in output, 3 in input and delay 0.5 seconds), no improvement could be obtained. This means that the real structure of the system better corresponds to the ARX model than to the ARMAX model, or that there is not enough data available to estimate the extra parameters.

3) Output Error: An OE model has the following form:

$$y(t) \quad = \quad \frac{B(q)}{F(q)}u(t-d) + e(t) \tag{5}$$

In the Output Error case the best results (making the trade-off between results and model complexity) are also obtained for input order 3 and output order 0 with a delay of 0.5 seconds ($oe6$) (see Table II).

TABLE II

RESULTS OF OE WITH DIFFERENT ORDERS AND DELAYS, 3 SECONDS AHEAD PREDICTION

ID	Order B	Order F	Delay(s)	Perf. $20Hz$	Perf. $5Hz$
oe1	5	5	0	77.2%	78.1%
oe2	1	5	0	59.2%	79.1%
oe3	5	0	0	69.8%	78.7%
oe4	5	0	0.5	78.0%	79.1%
oe5	4	0	0.5	77.5%	79.0%
oe6	3	0	0.5	75.5%	77.7%
oe7	2	0	0.5	62.9%	65.7%
oe8	3	0	0	51.8%	63.4%

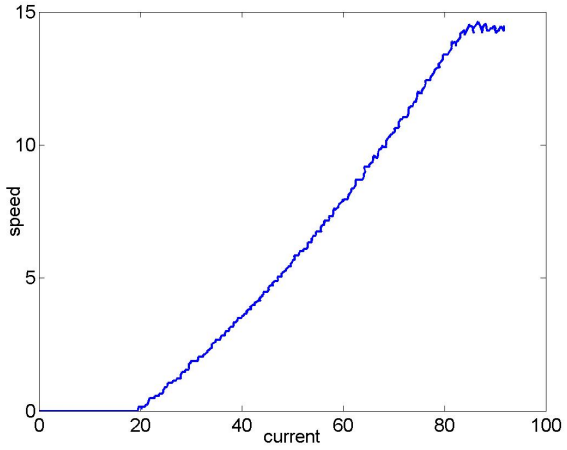


Fig. 4. Response in speed to a slope on the input current

4) Conclusion: The ARX and OE model structures deliver similar results. It may in the future however be necessary to update the model online to take different soils into account. In this view ARX is preferred because it can easily be implemented recursively [8] [9].

C. Non-linear preprocessing

A slope was applied to the input current in order to check for non-linearities in the system (see Fig. 4). This experiment was performed twice and gave practically identical results. There is a dead zone visible for small and large currents. In between there seems to be a quadratic relationship between input current and ground speed.

The relevance for model performance is shown in Table III. The best performing ARX models of Table I ($arx7$, $arx10$) (A) are compared to a model with the same parameters (order 3 in the input and order 0/1 in output) but fitted on transformed input data (B).

D. Non-linear models

As mentioned before, if the signals are filtered at more than $1Hz$, linear modelling becomes very hard. This gives reason to believe that the bump in the FRF between $1Hz$ and $2Hz$ (see Fig. 2) is caused by a nonlinearity.

In recent literature [10] [11] [12] Hammerstein-Wiener modelling is often used to introduce some non-linearity in a

TABLE III

RESULTS OF ARX ON CURRENT (A) AND QUADRATICALLY TRANSFORMED CURRENT (B), 3 SECONDS AHEAD PREDICTION

ID		# A	# B	Delay(s)	Perf. $20Hz$	Perf. $5Hz$
arx7	A	0	3	0.5	75.1%	76.4%
arx7B	B	0	3	0.5	79.1%	81.8%
arx10	A	1	3	0.5	68.2%	72.3%
arx10B	B	1	3	0.5	75.8%	79.1%

#: order

TABLE IV

RESULTS OF NONLINEAR ARX IDENTIFICATION ON CURRENT, 3 SECONDS AHEAD PREDICTION

ID	#	# A	# B	Delay(s)	Perf. $20Hz$	Perf. $5Hz$
narx1	2	2	3	0.0	81.0%	80.8%
narx2	3	2	3	0.0	78.7%	79.1%

#: degree or order
- indicates an unstable prediction model

linear dynamic model. In this section a second and third order input nonlinearity (Hammerstein) is considered. As dynamic system the ARX system is taken. The input nonlinearity $f(u)$ is expressed as follows:

$$f(u) = aH_1(u) + bH_2(u) + cH_3(u) \quad (6)$$

with $H_n(u)$ the Hermite polynomial [13] of order n. This expression for the non-linearity is used because the Hermite polynomials are orthogonal basis functions with respect to e^{-x^2} on the interval $[-\infty, +\infty]$ [14].

a) Condition: The problem sketched above is badly conditioned because of the input and output filtering. Due to low-pass filtering, sequential input/output measurements will be correlated, which increases the correlation between columns in the block-Hankel matrices, and increases the condition number of the latter matrix. This in turn reflects negatively on the variance on the estimates of the obtained least squares solution [15]. This bad condition can be resolved by downsampling the data, or by introducing a regularization term [16]. This last option is chosen. This means that the equation $Ax = B$ is solved in the following way:

$$x = (A^T A + \gamma I)^{-1} A^T B \quad (7)$$

with γ a regularization parameter and I the unity matrix.

b) Models: The best results are obtained with a model with order 2 in the output, order 3 in the input and delay 0. In Table IV the results are shown for a second degree and a third degree Hermite polynomial in the input current.

These models give analogous results to linear ARX models. Taking the number of parameters in these nonlinear ARX models into account however, linear ARX (possibly with nonlinear preprocessing) should be preferred.

E. Prediction of step response

One of the most common control inputs is a step. The best performing models are now validated on such an input sequence. It immediately becomes clear that output feedback of some sort is needed in order to improve model performance on this type of control inputs. This type of model performs a bit worse on the full noise data, but is much more robust to model imperfections. Moreover, the ARX model with output feedback ($arx10$) (with and without preprocessing) is the only model so far with practically white residuals.

The response of $arx7$, $arx7B$, $arx10$ and $arx10B$ is shown together with the real machine response in Fig. 5. It is clear that output feedback combined with nonlinear input preprocessing offers considerable advantages for the 3 seconds ahead prediction of input steps (see Table V).

TABLE V

RESULTS ON STEP INPUT, 3 SECONDS AHEAD PREDICTION

ID	Perf. $20Hz$
arx7	51.4%
arx10	23.5%
arx7B	58.8%
arx10B	91.3%
narx1	61.3%
narx2	-

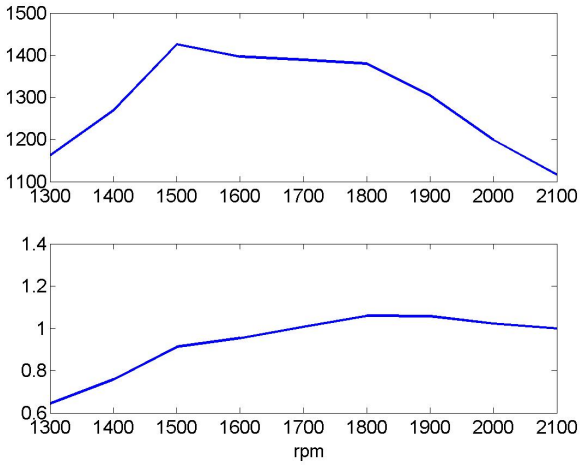


Fig. 6. Torque (above) and relative power (below) in function of rpm level

V. EXTENDED ROAD MODEL

A. Different RPM levels

So far all the experiments were performed at maximum rpm level of the diesel engine. In practice the rpm level can vary between 1300 rpm and 2100 rpm. The dynamic response of the machine depends of course on the rpm level.

The torque also varies as a function of the rpm level. A valid assumption might be that the dynamic behavior is only scaled with the available power. In Fig. 6 the maximum power delivered by the diesel engine in function of rpm level is shown (scaled to 1 at 2100 rpm). By scaling the predicted ground speed with this factor, the model remains valid for other rpm levels.

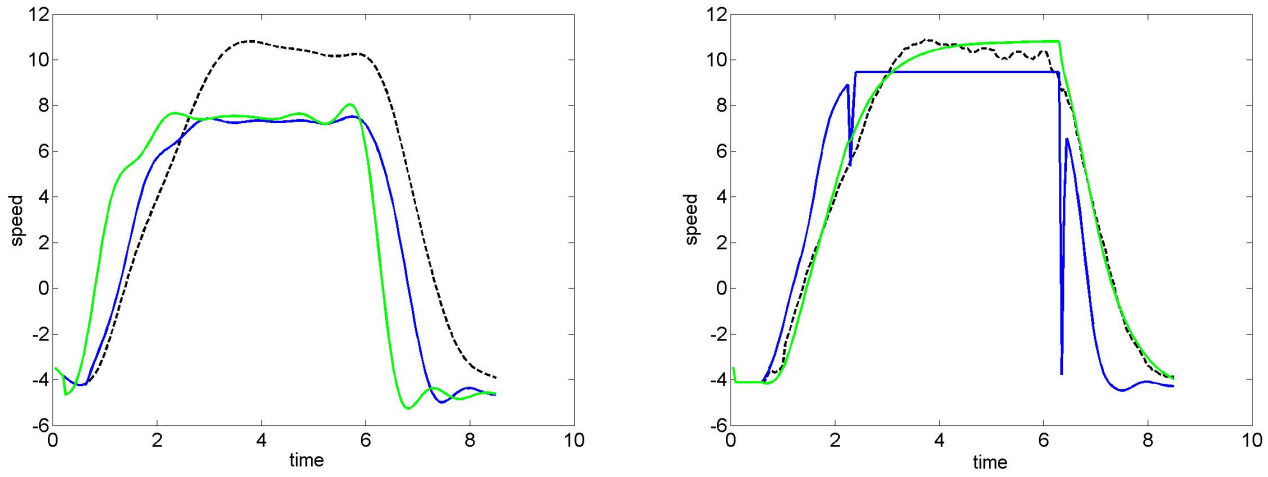


Fig. 5. Step response (3 seconds ahead prediction) of $arx7$ (blue) and $arx10$ (green) (left) and of $arx7B$ (blue) and $arx10B$ (green) (right) together with the measured output (black dashed)

TABLE VI
RELATIVE DC GAINS FOR DIFFERENT SLOPES AND MASSES

	Slope	Mass	DC gain(dB)	DC gain(%)
Downhill	-2^o	15700	0.05	100.6%
	-3.5^o	17200	0.39	104.6%
	-2^o	18800	0.42	105.0%
Flat		15700	0	100%
		17200	-0.23	97.4%
		18800	-0.25	97.2%
Uphill	$+3.3^o$	15700	-0.92	90.0%
	$+4.4^o$	17200	-1.16	87.5%
	$+2.8^o$	18800	-0.87	90.5%

B. Influence of mass and slope

The step response on slopes and with different masses was practically identical. However the machine FRF (obtained by using white noise as control input) shows a clear influence of mass and slope in the DC area. At higher frequencies the difference is too small compared to the noise to draw conclusions.

The DC gain of the different tests relative to the test with the default mass on flat terrain is shown in Table VI, together with the corresponding mass and slope. These spectra are derived on quite short noise sequences of about 1 minute, and should thus be interpreted with the necessary care. A couple of trends are however clear from this data. The machine runs faster downhill and slower uphill, depending on the slope and the vehicle mass. On flat terrain, a heavier machine runs slower than a lighter one.

Intuitively, a machine is slowed down by the friction force F_F and slowed down or accelerated by the gravity force F_G:

$$F_F = \beta m g \quad (8)$$

$$F_G = m g \sin(\alpha) \quad (9)$$

with β the friction constant, m the total mass, g the gravitational acceleration and α the vehicle slope. These effects are also visible in the FRFs.

In the step responses of the system however, on different slopes and with different masses, this influence was not clear. This is because the effect of mass and slope on the road is not large enough for the powerful diesel engine. In field conditions, where the engine faces higher loads, the mass and slope effects will play a more important role.

C. Different gears

All the experiments described above were performed in 2^{nd} gear. This is the gear commonly used during harvesting operation. In this section the model designed on 2^{nd} gear data is applied to the other gears.

The time to reach maximum speed in 1^{st} gear is the same as the time do so in 2^{nd} gear. This means that rescaling the model for different maximum speeds (before subtracting the mean) should suffice to obtain good results.

For 3^{rd} and 4^{th} gear the reduction ratio of the mechanical transmission has to be taken into account. To adapt the model, the torque is assumed to be proportional to the change of speed Δv. This means that, based on the preservation of power, the following equation should hold:

$$v_{max,3}\Delta v_3 = v_{max,2}\Delta v_2 \quad (10)$$

with $v_{max,i}$ the maximum speed reached in i-th gear, and Δv_i the acceleration in i-th gear. With equation (3), the acceleration for an ARX model $y = f(u, y)$ can be rescaled as follows:

$$\Delta v = y_k - y_{k-1} = \alpha(f(u, y) - y_{k-1}) \quad (11)$$

$$y_k = \alpha f(u, y) + (1 - \alpha) y_{k-1} \quad (12)$$

The model can then be rescaled for other gears by choosing α such that (10) is satisfied. The results of this extra rescaling, after adding the mean again, is shown in Fig. 7.

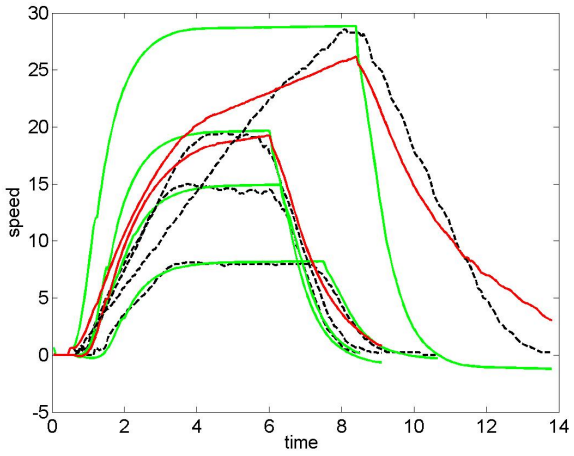


Fig. 7. Step input (3 seconds ahead prediction) on the 4 gears modelled by rescaled $arx10B$ (green) and acceleration rescaled $arx10B$ (red) together with the measured output (black dashed)

VI. CONCLUSIONS AND FUTURE WORK

A. Conclusions

In this paper a model for the 2^{nd} gear of a New Holland CR combine harvester was developed. Nonlinear input preprocessing in combination with a linear ARX model ($arx10B$) delivered the best results. The model was validated on an independent white noise sequence as well as on input steps.

The model was then extended to incorporate different RPM levels. This was done based on power considerations. The influence of machine mass and road slope was also examined. In road conditions this did not influence the machine dynamics significantly. In field conditions however, mass and slope are expected to gain importance.

Finally the performance of the model, after rescaling of maximum value and acceleration, in the other gears is illustrated on step inputs. The model performs very well in 1^{st}, 2^{nd} and 3^{rd} gear, and acceptable in 4^{th} gear.

B. Future Work

A combine operates in the field. These are completely different circumstances than a concrete road. The model derived here will have to be validated in the field. Since field conditions have significant influence on the damping of the machine, it may be necessary to make the model adaptive.

Mass and slope influence will become significant in field conditions, since the diesel engine then faces a much larger load. Validation in the field thus will also mean including a model for mass and slope influence.

For other gears than 2^{nd}, the machine model still has to be validated on a wider range of input sequences.

ACKNOWLEDGMENTS

T. Coen is currently supported by a scholarship of I.W.T.-Vlaanderen (Institute for the Promotion of Innovation through Science and Technology in Flanders). I. Goethals is an assistant researcher with the FWO-Vlaanderen. J. Anthonis is a postdoctoral researcher with the FWO-Vlaanderen. B. De Moor and J. De Baerdemaeker are full professors with the Katholieke Universiteit Leuven, Belgium.

The author gratefully acknowledges the cooperation of CNH Belgium N.V..

REFERENCES

[1] H. Kutzbach, "Trends in power and machinery," *J. agric. Engng Res.*, vol. 76, pp. 237–247, 2000.

[2] K. Maertens, H. Ramon, and J. De Baerdemaeker, "An on-the-go monitoring algorithm for separation processes in combine harvesters," *Computers and electronics in agriculture*, vol. 43, pp. 197–207, 2004.

[3] J. Anthonis, K. Maertens, G. Strubbe, J. De Baerdemaeker, and H. Ramon, "Design of a friction independant mass flow sensor by force measurement on a circular chute," *Biosystems Engineering*, vol. 84, no. 2, pp. 127–136, 2003.

[4] K. Maertens and J. De Baerdemaeker, "Flow rate based predicition of threshing process in combine harvesters," *Applied Engineering in Agriculture*, vol. 19, no. 4, pp. 383–388, 2003.

[5] A. Vahidi, A. Stefanopoulou, and H. Peng, "Recursive least squares with forgetting for online estimation of vehicle mass and road grade: Theory and experiments," *Vehicle System Dynamics*, submitted 2003.

[6] M. Jelali and A. Kroll, *Hydraulic Servo-systems: Modelling, Identification and Control*. Springer-Verlag, Berlin, Germany, 2003.

[7] L. Ljung, "System identification toolbox user's guide," The Mathworks, Inc., 2000.

[8] V. Smidl, A. Quinn, M. Karny, and T. Guy, "Robust estimation of autoregressive processes using a mixture-based filter-bank," *Systems & Control Letters*, in press.

[9] B. Lindoff, "On parameter estimation and control of time-varying stochastic systems," PhD thesis, Dept. Mathematical Statistics, Lund Institute of Technology, Lund, Sweden, 1997.

[10] Y. Zhu, "Estimation of an n-l-n hammerstein-wiener model," *Automatica*, vol. 38, pp. 1607–1614, 2002.

[11] G. Dolanc and S. Strmcnik, "Identification of nonlinear systems using a piecewise-linear hammerstein model," *System & Control Letters*, vol. 54, pp. 145–158, 2005.

[12] E. Bai, "A blind approach to the hammerstein-wiener model identification," *Automatica*, vol. 38, pp. 967–979, 2002.

[13] C. Hermite, "Sur un nouveau développement en série de fonctions," *Compt. Rend. Acad. Sci. Paris*, vol. 58, pp. 93–100, 266–273, 1864.

[14] N. Lebedev, *Special Functions & their applications*. Dover Publications, England, 1972.

[15] G. Golub and C. Van Loan, *Matrix Computations*. John Hopkins University Press, 1989.

[16] O. Nelles, *Nonlinear System Identification*. Springer-Verlag, Berlin, Germany, 2001.