CropGym: a reinforcement learning environment for crop management

Hiske Overweg,[*] **Herman N. C. Berghuijs, Ioannis N. Athanasiadis**
Wageningen University and Research, The Netherlands

Abstract

Nitrogen fertilizers have a detrimental effect on the environment, which can be reduced by optimizing fertilizer management strategies. We implement an OpenAI Gym environment where a reinforcement learning agent can learn fertilization management policies using process-based crop growth models and identify policies with reduced environmental impact. In our environment, an agent trained with the Proximal Policy Optimization algorithm is more successful at reducing environmental impacts than the other baseline agents we present.

1 Introduction

Fertilizer use has vastly improved crop yields in the past century (Frink et al., 1999). Yet excessive use of fertilizers is damaging the environment in several ways (Schlesinger, 2009; Sheriff, 2005; Yadav et al., 1997). Leakage into waterways can lead to eutrophication and lower oxygen levels up to an extent where marine life is no longer supported (Diaz & Rosenberg, 2008). On top of that, excessive fertilization contributes to global warming due to high energy demand of fertilizer production and conversion by soil bacteria into nitrous oxide, a potent green house gas (Erisman et al., 2011).

By improved farming practises, leakage into the environment can be reduced. As an example, improvements can come from tailored side dress recommendations, where at the start of the season a small dose of fertilizer is applied which gets supplemented locally during the growth season (van Evert et al., 2012). Alternatively, intercropping methods (Cong et al., 2015) and other eco-innovations can reduce the need for synthetic fertilization (Hasler, 2017).

Studying and validating farming strategies in practise is a challenging process, because of the long duration and labor intensity of field trials and the presence of many confounding factors. Process-based models, which simulate crop growth, can help to gain insights in this area. In this work, we create a reinforcement learning environment in which an agent can learn fertilization strategies using a process-based crop growth model. While reinforcement learning has had significant impact in other domains, we think its potential in agriculture has not been fully recognized yet. By providing an OpenAI Gym environment, we aim to encourage exploration of the application of reinforcement learning to sustainable agricultural management.

2 Related work

Deep RL is the subfield of machine learning which is concerned with training agents to take actions in an environment based on a reward signal (Sutton, 2018). Deep RL has been successfully applied in numerous fields, ranging from computer games (Silver et al., 2016; Mnih et al., 2013), to the autonomous navigation of stratospheric balloons (Bellemare et al., 2020). Applications to agriculture have been proposed in Bu & Wang (2019) and Binas et al. (2019). RL agents have been implemented to control irrigation in a greenhouse setting (Zhou, 2020) and to determine field sampling strategies using remote sensing (Zhang et al., 2020). We are not aware of any works which study crop fertilization strategies or make use of process-based crop growth models.

OpenAI Gym provides a standardized collection of benchmark problems with a common interface (Brockman et al., 2016). It allows for comparison between algorithms and assessment of generaliza-

---

[*]Corresponding author: hiske.overweg@wur.nl

tion performance. It also provides an interface to convert a real-world problem into a custom Gym environment.

Process-based crop growth models are designed for quantitative analysis of biophysical processes in crop growth and production. These models describe crop growth processes in terms of differential equations which estimate crop growth in daily timesteps, based on factors such as light interception, water and nutrient availability. Widely used frameworks are APSIM (Keating, 2003; Holzworth, 2014) and the Python Crop Simulation Environment (PCSE), which contains models as LINTUL-3 (Shibu et al., 2010) and WOFOST (WOrld FOod STudies, de Wit et al. (2019)). For a literature review of agricultural systems modelling, see Jones (2017). Modern wheat crop models are discussed in Chenu et al. (2017).

3 DESCRIPTION OF THE GYM ENVIRONMENT

The CropGym environment[1] contains the following components:

- The state space S, consisting of the current state of the crop (in terms of a multidimensional output of a process-based crop growth model) and an multidimensional weather observation
- The action space A, which consists of discrete doses of fertilizer to apply
- Transitions between states, governed by the deterministic process-based model and the weather sequence
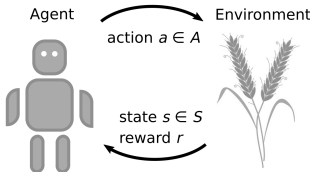- The reward r, which encourages large yield and limited fertilizer use.

Figure 1: Interaction between the agent and the CropGym environment

The information flow is shown schematically in Fig. 1. The above framework can be applied for any process-based crop growth model or farm management simulation. In the next sections we detail our implementation choices.

State space A natural intervention interval for farmers is a week. This is why, using the default settings, the agent observes and intervenes in the CropGym environment on a weekly basis. The observations by the agent consist of the output variables of the process-based crop growth model LINTUL-3 (Light INtercepion and UTilization). Nitrogen limited crop growth in LINTUL-3 is implemented in the Python Crop Simution Environment (PCSE). The model parameters have been calibrated to simulate winter wheat (Wiertsema, 2015), which is why we chose to focus on this crop. The observed variables from the crop growth model are listed in Appendix A.1. Additionally the agent observes the weather of the past week as described by the default weather variables in PCSE (see Appendix A.2). For these weather variables we use 29 years of weather data from 9 locations in the Netherlands from the PowerNASA database.

Action space The agent's action space consists of a range of nitrogen fertilizer values to apply at each time step. From a farm management perspective, applying no fertilizer at all at a given point in time is an attractive choice, because there is no time investment. This is why we chose to discretize the action space. At each intervention, the agent has the following options for fertilizer application:

$$A = \{20\,k\tfrac{\text{kg}}{\text{ha}} \mid k \in \{0, 1, 2, ..., 6\}\}$$

In this way the agent can experiment with both doses which are smaller than and similar to those used in practise (Wiertsema, 2015).

---

[1]The source code is available at **link to be added in camera-ready version**

Reward We reward the agent for a large grain yield (we do not take losses during harvest into account), which is referred to in LINTUL-3 as the mass of the storage organ (m_{SO}) of the crop, expressed in kilograms of dry matter per hectare. The achievable mass varies per season, because weather variables affect the yield. To account for this, we calculate both the mass m_{SO} the agent achieves with its fertilization policy, and the mass m_{SO}^* which would have been reached without fertilizer application. Our reward function contains the difference between these quantities. Finally, our goal is to discourage fertilizer application, which is why we penalize the agent for the weight of fertilizer applied (m_{fert}), expressed in the same units. The reward at timestep t is thus defined as:

$$r_t = m_{SO,t} - m_{SO,t-1} - (m_{SO,t}^* - m_{SO,t-1}^*) - \beta m_{fert,t}$$

where the parameter β determines the trade-off between large yield and reduced environmental impact. Setting $\beta \sim 1$ leads to optimization of the current economic cost for a farmer in the Netherlands, since wheat prices and fertilizer prices are similar (agrimatie.nl, a;b). When developing policies which have reduced environmental impact, we need to set $\beta > 1$.

4 BASELINE AGENTS

Implemented agents To determine whether reinforcement learning agents can indeed identify efficient fertilization policies, we compare three agents:

- The standard practise agent adds applies nitrogen fertilizer on three different dates during the season, as is done in practise in the Netherlands. For details, see Appendix A.3.
- The reactive agent adds a fixed amount of fertilizer to the field whenever the nitrogen in the soil is depleted.
- The PPO agent is trained with the Proximal Policy Optimization algorithm, a widely used policy gradient algorithm (Schulman et al., 2017), as implemented in the Stable Baselines3 package (Raffin et al., 2019). The policy and value function networks are both fully connected networks consisting of 2 layers with 64 nodes per layer.

For more details on agent training and hyperparameters, see Appendix A.3.

Agent performance We evaluate the reward the agents obtain on the seasons starting in 1984, 1994, 2004 and 2014 with weather data from 52^oN, 5.5^oE. These years were excluded from the training set of the PPO agent. For the reward function we set $\beta = 10$. As can be seen in Fig. 2a, the PPO agent gets a higher reward than the standard practise agent and the reactive agent in 1994, 2004 and 2014. In Fig. 2b,c we show the total yield obtained and the amounts of fertilizer applied by the agents respectively. The PPO agent has a somewhat smaller yield than the other agents in 1984 and 2004, because of the reduced fertilization. Figure 2d shows the soil nitrogen level and the fertilization events proposed by the PPO agent in 2014. The nitrogen uptake by the wheat and the reward are shown in Fig. 2e,f respectively. Similar data for other years from the test set is shown in Appendix A.4.

5 CONCLUSION

We introduce a new Gym environment to study fertilization strategies for crops and discuss the considerations we made during the design process of the CropGym. We implement a reactive agent, standard practise and an agent trained with PPO, which serve as a baseline for further investigations. The PPO agent was able to learn strategies with a higher reward than the other agents for three out of four years in the test set, which demonstrates that reinforcement learning agents are able to identify policies with reduced environmental impact.

6 FUTURE WORK

Other farm management practices After further calibration of the more detailed process-based model WOFOST, we intend to train an agent to learn policies for more management options, as phosphorus and potassium fertilization. Learning irrigation policies and including weather forecast
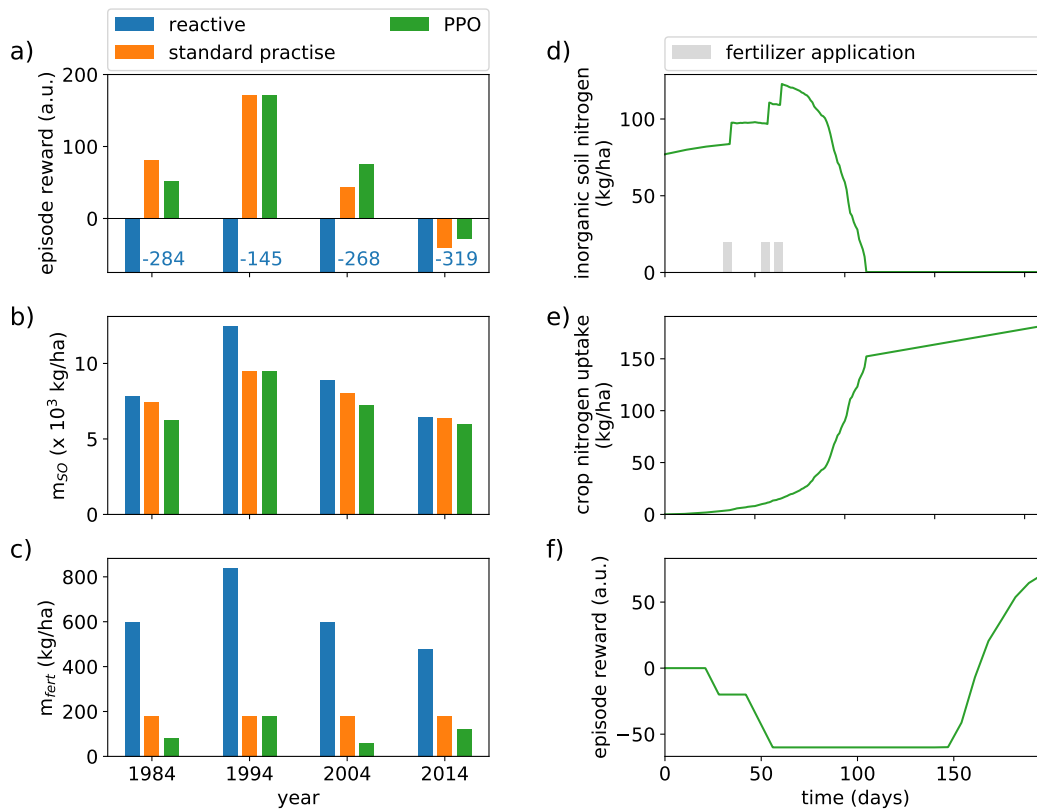
Figure 2: Agent performance. (a) Reward obtained by various agents. (b) The weight of the storage organ at the end of the season. (c) The amount of fertilizer applied by the agents. (d) Nitrogen available in the soil as a function of time in the season starting in 2004. Gray bars indicate fertilization events by the PPO agent. (e) Wheat nitrogen uptake. (f) Reward of the PPO agent.

data are other extensions we consider. The potential of reinforcement learning agents can be further exploited at a large scale when using whole farm simulations, which take interactions between fields and livestock farming into account (Holzworth, 2014; Rodriguez et al., 2014).

Environmental impact and climate change When coupling process-based crop growth models to more detailed soil models, the interaction between fertilization, rainfall and irrigation can be taken into account (Engel & Priesack, 1993; Groenendijk et al., 2016). Just as process-based models, the CropGym can be used to study the effect of a changing climate on crop growth, to assess adaptations of farm management strategies (Rodriguez et al., 2014).

Reality gap To get closer to the practical situation in the field, the environment can be modified to a partially observable Markov Decision Process, in which the agent only observes variables which can be realistically observed on a weekly basis. More accurate digital representations of crop growth processes would also help to get closer to reality. These can be provided by tailored calibration of process-based crop growth models, inclusion of other experimental data sources within the models (Holzworth, 2014) or digital farm twins (Jans-Singh et al., 2020; Pylianidis et al., 2021).

REFERENCES

agrimatie.nl. Price development of seeds and grains https://www.agrimatie.nl/SectorResultaat.aspx?subpubID=2232&sectorID=2233&themaID=2263. a. Accessed: 2021-03-09.

agrimatie.nl. Price development of fertilizer https://www.agrimatie.nl/ThemaResultaat.aspx?subpubID=2289&themaID=2263. b. Accessed: 2021-03-09.

Marc G. Bellemare, Salvatore Candido, Pablo Samuel Castro, Jun Gong, Marlos C. Machado, Subhodeep Moitra, Sameera S. Ponda, and Ziyu Wang. Autonomous navigation of stratospheric balloons using reinforcement learning. *Nature*, 588(7836):77–82, 2020.

Jonathan Binas, Leonie Luginbuehl, and Yoshua Bengio. Reinforcement Learning for Sustainable Agriculture. *CCAI workshop at the 36 th International Conference on Machine Learning*, 2019.

Greg Brockman, Vicki Cheung, Ludwig Pettersson, Jonas Schneider, John Schulman, Jie Tang, and Wojciech Zaremba. OpenAI Gym. pp. 1–4, 2016. URL http://arxiv.org/abs/1606.01540.

Fanyu Bu and Xin Wang. A smart agriculture IoT system based on deep reinforcement learning. *Future Generation Computer Systems*, 99:500–507, 2019.

Karine Chenu, John Roy Porter, Pierre Martre, Bruno Basso, Scott Cameron Chapman, Frank Ewert, Marco Bindi, and Senthold Asseng. Contribution of Crop Models to Adaptation in Wheat. *Trends in Plant Science*, 22(6):472–490, 2017.

Wen Feng Cong, Ellis Hoffland, Long Li, Johan Six, Jian Hao Sun, Xing Guo Bao, Fu Suo Zhang, and Wopke Van Der Werf. Intercropping enhances soil carbon and nitrogen. *Global Change Biology*, 21(4):1715–1726, 2015.

Allard de Wit, Hendrik Boogaard, Davide Fumagalli, Sander Janssen, Rob Knapen, Daniel van Kraalingen, Iwan Supit, Raymond van der Wijngaart, and Kees van Diepen. 25 years of the WOFOST cropping systems model. *Agricultural Systems*, 168(June 2018):154–167, 2019.

Robert J. Diaz and Rutger Rosenberg. Spreading dead zones and consequences for marine ecosystems. *Science*, 321(5891):926–929, 2008.

Thomas Engel and Eckart Priesack. Expert-n-a building block system of nitrogen models as resource for advice, research, water management and policy. In *Integrated soil and sediment research: A basis for proper protection*, pp. 503–507. Springer, 1993.

Jan Willem Erisman, Jim Galloway, Sybil Seitzinger, Albert Bleeker, and Klaus Butterbach-Bahl. Reactive nitrogen in the environment and its effect on climate change. *Current Opinion in Environmental Sustainability*, 3(5):281–290, 2011.

Charles R. Frink, Paul E. Waggoner, and Jesse H. Ausubel. Nitrogen fertilizer: Retrospect and prospect. *Proceedings of the National Academy of Sciences of the United States of America*, 96(4):1175–1180, 1999.

Piet Groenendijk, Hendrik Boogaard, Marius Heinen, JG Kroes, Iwan Supit, and Allard de Wit. Simulation nitrogen-limited crop growth with swap/wofost: process descriptions and user manual. Technical report, Wageningen Environmental Research, 2016.

Kathrin Hasler. *Environmental impact of mineral fertilizers: possible improvements through the adoption of eco-innovations*. PhD thesis, Wageningen University, 2017.

Dean P. et al. Holzworth. APSIM - Evolution towards a new generation of agricultural systems simulation. *Environmental Modelling and Software*, 62:327–350, 2014.

Melanie Jans-Singh, Kathryn Leeming, Ruchi Choudhary, and Mark Girolami. Digital twin of an urban-integrated hydroponic farm. *Data-Centric Engineering*, 1, 2020.

James W. et al. Jones. Brief history of agricultural systems modeling. *Agricultural Systems*, 155: 240–254, 2017.

B. A. et al. Keating. An overview of APSIM, a model designed for farming systems simulation. *European Journal of Agronomy*, 18(3-4):267–288, 2003. doi: 10.1016/S1161-0301(02)00108-9.

Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller. Playing Atari with Deep Reinforcement Learning. pp. 1–9, 2013.

Christos Pylianidis, Sjoukje Osinga, and Ioannis N Athanasiadis. Introducing digital twins to agriculture. *Computers and Electronics in Agriculture*, 184(December 2020):105942, 2021.

Antonin Raffin, Ashley Hill, Maximilian Ernestus, Adam Gleave, Anssi Kanervisto, and Noah Dormann. Stable baselines3. `https://github.com/DLR-RM/stable-baselines3`, 2019.

Daniel Rodriguez, Howard Cox, Peter DeVoil, and Brendan Power. A participatory whole farm modelling approach to understand impacts and increase preparedness to climate change in Australia. *Agricultural Systems*, 126:50–61, 2014.

William H. Schlesinger. On the fate of anthropogenic nitrogen. *Proceedings of the National Academy of Sciences of the United States of America*, 106(1):203–208, 2009.

John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv*, pp. 1–12, 2017.

Glenn Sheriff. Efficient waste? Why farmers over-apply nutrients and the implications for policy design. *Review of Agricultural Economics*, 27(4):542–557, 2005.

M. E. Shibu, P. A. Leffelaar, H. van Keulen, and P. K. Aggarwal. LINTUL3, a simulation model for nitrogen-limited situations: Application to rice. *European Journal of Agronomy*, 32(4):255–271, 2010.

D Silver, J Schrittwieser, K Simonyan, I Antonoglou Nature, and Undefined 2017. Mastering the game of Go without human knowledge. *Nature*, 550(7676):354, 2016.

Richard S. Sutton. *Reinforcement Learning an introduction*. 2018. ISBN 9780262039246.

Frits K. van Evert, Remmie Booij, Jan Nammen Jukema, Hein F.M. ten Berge, Dik Uenk, E. J.J.Bert Meurs, Willem C.A. van Geel, Klaas H. Wijnholds, and J. J.Hanja Slabbekoorn. Using crop reflectance to determine sidedress N rate in potato saves N and maintains yield. *European Journal of Agronomy*, 43:58–67, 2012.

Wiert Wiertsema. Obtaining winter wheat parameters for LINTUL from a field experiment. 2015.

Satya N. Yadav, Willis Peterson, and K. William Easter. Do farmers overuse nitrogen fertilizer to the detriment of the environment? *Environmental and Resource Economics*, 9(3):323–340, 1997.

Zichen Zhang, Jayson Boubin, Christopher Stewart, and Sami Khanal. Whole-field reinforcement learning: A fully autonomous aerial scouting method for precision agriculture. *Sensors*, 20(22): 1–16, 2020.

Ni Zhou. Intelligent control of agricultural irrigation based on reinforcement learning. *Journal of Physics: Conference Series*, 1601(5), 2020.

A APPENDIX

A.1 CROP GROWTH VARIABLES

Symbol	Meaning	Unit
DVS	development stage	-
LAI	Leave area index	-
TGROWTH	Total biomass growth	g/m^2
NUPTT	Total N uptake	$g\ N/m^2$
TRAN	Crop transpiration rate	$m^3\ H_2O/m^2$
TIRRIG	Water applied by irrigation	$m^3\ H_2O/m^2$
TNSOIL	Soil nitrogen available	$g\ N/m^2$
TRAIN	Total rainfall	$m^3\ H_2O/m^2$
TRANRF	Transpiration reduction factor	mm/day
TRUNOF	Soil runoff	$m^3\ H_2O/m^2$
TAGBM	Total above ground dry weight	g/m^2
TTRAN	Water removed by transpiration	$m^3\ H_2O/m^2$
WC	Soil moisture content	$m^3\ H_2O/m^2$
WLVD	Dry weight of dead leaves	g/m^2
WRT	Dry weight of roots	g/m^2
WLVG	Weight green leaves	g/m^2
WSO	Dry weight of storage organ	g/m^2
WST	Dry weight of stems	g/m^2

A.2 WEATHER VARIABLES

Symbol	Meaning	Unit
TMAX	Daily maximum temperature	^oC
TMIN	Daily minimum temperature	^oC
VAP	Mean daily vapour pressure	hPa
RAIN	Precipitation	cm/day
IRRAD	Daily global radiation	$J/m^2/day$

A.3 AGENT HYPERPARAMETERS

Standard practise agent   The standard practise agent applied nitrogen on the following days, in correspondence with Wiertsema (2015):

Date	Amount of fertilizer
3rd of March	60 kg/ha
31st of March	60 kg/ha
5th of May	60 kg/ha

Reactive agent   The reactive agent adds 120 kg/ha of nitrogen fertilizer whenever the soil nitrogen drops below 5 kg/ha.

PPO agent   The environment was normalized with the VecNormalize environment wrapper, with a discount factor of γ=1, a normalized reward and observation clipping set to 10. The PPO agent trained for 500.000 timesteps using default hyperparameters.
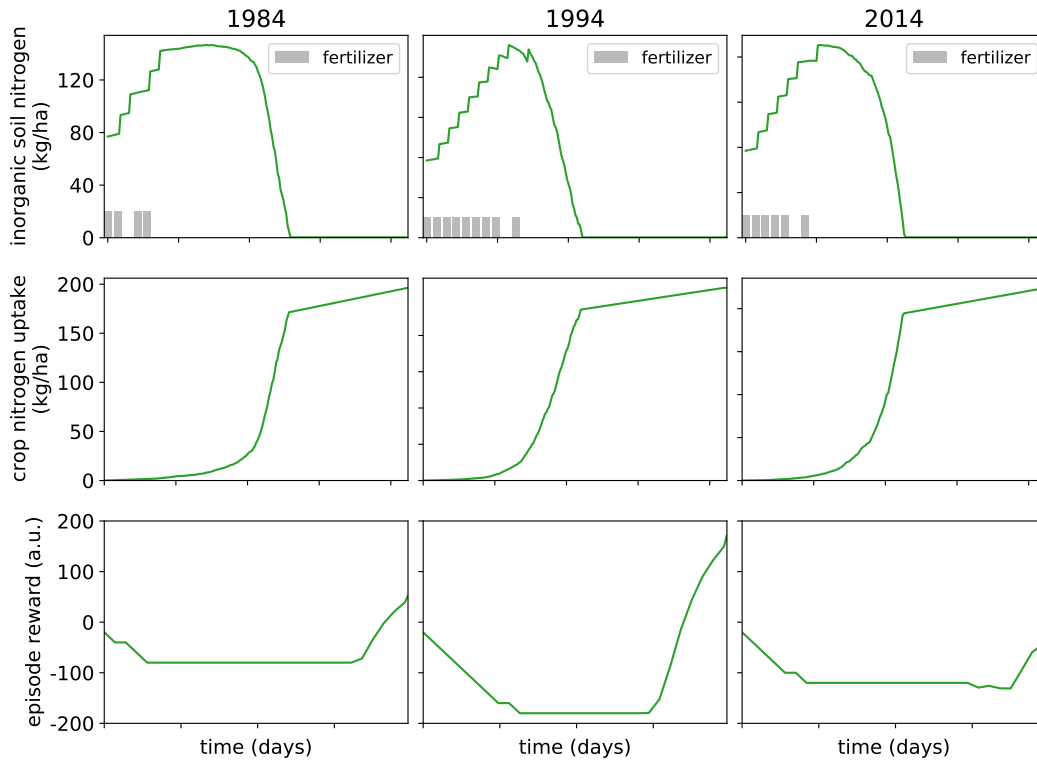
A.4 AGENT RESULTS

Figure A.1: Performance of the PPO agent in 1984, 1994 and 2014